21

Data Mining Method for Energy System Aplications

Reşat Selbaş, Arzu Şencan and Ecir U. Küçüksille
Department of Mechanical Education, Technical Education Faculty,
Süleyman Demirel University,
Turkey

1. Introduction

In recent years, data storage capacity of computers increased as well as their speeds. With these advances in the technology, millions of data can be recorded in computer memories. These data can be used for a better service to customers by companies. For example, millions of people are shopping in supermarkets everyday. These shopping data are added to companies' records every day. Information about customers' shopping habits can be obtainede by the analysis of these data and more efficient service can be offered by companies. However, companies don't use these data efficiently. Therefore, these data stay just as records in computer memories.

"Data mining process" is a term for efficient usage of data. There are many different definitions for data mining process in literature. Data mining is a step in the KDD process that consists of applying data analysis and discovery algorithms that produce a particular enumeration of patterns (or models) over the data [1]. Data mining is an interdisciplinary field bringing together techniques from machine learning, pattern recognition, statistics, databases, and visualization to address the issue of information extraction from large databases [2]. This technology is motivated by the need of new techniques to help analyze, understand or even visualize the huge amounts of stored data gathered from business and scientific applications. It is the process of discovering interesting knowledge, such as patterns, associations, changes, anomalies and significant structures from large amounts of data stored in databases, data warehouses, or other information repositories [3].

Data mining process is used in finanel, health, communication, medicine and science fields. The most vivid example is "Amazon" web site.

Many computer programs are used for data mining processes. Some are open source and some cost money. While WEKA, Rapid Miner, Pentaho, Orange, Scriptella ETL(Extract-Transform-Load), KNIME, ELKI(Environment for DeveLoping KDD-Applications Supported by Index-Structures) programs are open source, SPSS Clementine, Sql Server Business Intelligent Studio, SAS Data Mining Software, ODM(Oracle Data Mining) Softwares cost money.

In this chapter, DM process used in the energy systems is reviewed. Available literature summaries published in this area is presented.

2. Data mining process

In Table 1, side-by-side comparison of the major existing KDDM models is shown [4].

Model	Fayad et al.	Cabena et al.	Anand&Buchner	CRISP-DM	Cios et al.	Generic Model
Area	Academic	Industrial	Academic	Industrial	Academic	N/A
No of Steps	9	5	8	6	6	6
Refs	(Fayyad et al., 1996d)	(et al., 1998)	(Anand & Buchner, 1998)	(Shearer, 2000)	(Cios et al., 2000)	N/A
Steps	1 Developing and Understanding of the Application Domain 2 Creating a Target Data Set 3 Data Cleaning and Preprocessing 4 Data Reduction and Projection 5 Choosing the DM Task 6 Choosing the DM Algorithm 7 DM 8 Interpreting Mined Patterns 9 Consolidating Discovered Knowledge	1 Business Objectives Determination 2 Data Preparation 3 DM 4 Domain Knowledge Elicitation 5 Assimilation of Knowledge	1 Human Resource Identification 2 Problem Specification 3 Data Prospecting 4 Domain Knowledge Elicitation 5 Methodology Identification 6 Data Preprocessing 7 Pattern Discovery 8 Knowledge Post-processing	1 Business Understanding 2 Data Understanding 3 Data Preparation 4 Modeling 5 Evaluation 6 Deployment	1 Understanding the Problem Domain 2 Understanding the Data 3 Preparation of the Data 4 DM 5 Evaluation of The Discovered Knowledge 6 Using the Discovered Knowledge	1 Application Domain Understanding 2 Data Understanding 3 Data Preparation and Identification of DM Technology 4 DM 5 Evaluation 6 Knowledge Consolidation and Deployment

Table 1. Side-by-side comparison of the major existing KDDM models

2.1 CRISP-DM data mining process

A systematic approach is essential to obtain satisfactory results for the DM analysis. Nowadays, a number of versions of DM tools exist. The most widespread application amongst the tools is CRISP-DM. CRISP-DM was developed by a consortium which consist the firms of NCR System Engineering (USA-Denmark), Daimler-Chrysler AG (Germany), SPSS Inc. (USA) and OHRA Verzekeringen en Bank Groep B.V (Netherlands) [5,6]. CRISP-DM is a process which defines the basic stages of DM, as can be seen from Fig. 1.

Business understanding

This initial phase focuses on understanding the project objectives and requirements from a business perspective, and then converting this knowledge into a data mining problem definition, and a preliminary project plan designed to achieve the objectives. [6, 7].

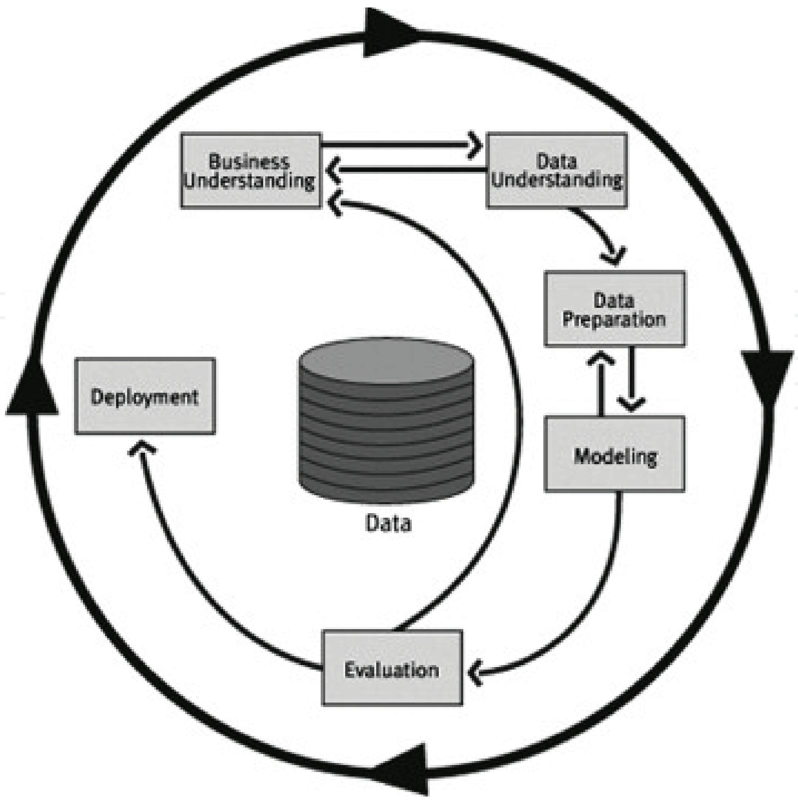

Fig. 1. CRISP-DM data mining process [5]

Data understanding

Second stage is called as the understanding of the data content. Data stored in a various environments such as Microsoft stores data in a hundred different database and 70 different data warehouses. First step is getting the meaningful data from those database or data warehouses for the selected application. In the meantime, in this phase data quality and discovering first insights into the data are seen as two important aspects.

Data preparation

The data preparation phase covers all activities to construct the final data set or the data that will be fed into the modeling tool from the initial raw data. Tasks include table, record, and attribute selection, as well as transformation and cleaning of data for modeling tools. The five steps in data preparation are the selection of data, the cleaning of data, the construction of data, the integration of data, and the formatting of data [8].
The purpose of the cleaning stage is selecting unsuitable or incorrectly entered data in the data. For example, filling the mean value for the instead of the incomplete data or erasing abnormal data records outside of the normal dispersion area assuming the meaningful data is in the normal distribution [6].
Data conversion is required for recording data in different formats or values since some data mining algorithms work only with data in digital format. In this case it needs to convert data in text format to the digital one. Purpose of the feature selection is determination of the most dominant parameters in forecasting a value. It might be assigned many features to estimate a value. However, it is not always easy to collect the determined data. For this case, by finding the effective properties data acquisition can be fast and simple. In addition, data are divided into two groups as training and testing data.

Modeling

In this phase, various modeling techniques are selected and applied and their parameters are calibrated to optimal values. Typically, several techniques exist for the same data mining problem type. Some techniques have specific requirements on the form of data. Therefore, stepping back to the data preparation phase may be necessary. Modeling steps include the selection of the modeling technique, the generation of test design, the creation of models, and the assessment of models [8]. If the task is fully accomplished, in this case, selection of the correct algorithm is much easier. Each task requires different algorithms and it is not known which one gives the best result without constructing the model. It may be only possible to guess according to the condition of the data in hand. If there is a linear relationship between whole input and estimation variables, choosing the decision tree algorithm can be good choice. If there is a complex relation among the variables, in this situation neural network algorithm can be selected [9]. Some of these algorithms are linear regression, multi layer perceptron, KStar, decision trees, K-means.

Evaluation

Before proceeding to final deployment of the model built by the data analyst, it is important to more thoroughly evaluate the model and review the model's construction to be certain it properly achieves the business objectives. Here it is critical to determine if some important business issue has not been sufficiently considered. At the end of this phase, the project leader then should decide exactly how to use the data mining results. The key steps here are the evaluation of results, the process review, and the determination of the next steps [8]. In this point, there are several tools exist. For instance, if there digital data exist for the estimation and wanted to test the accuracy of the model, RMSE (root mean square error) or R^2 (correlation coefficient) can be used [6].

Deployment

Model creation is generally not the end of the project. The knowledge gained must be organized and presented in a way that the customer can use it, which often involves applying "live" models within an organization's decision-making processes [8].
Depending on the requirements, the deployment phase can be as simple as generating a report or as complex as implementing a repeatable data mining process across the enterprise. Even though it is often the customer, not the data analyst, who carries out the deployment steps, it is important for the customer to understand up front what actions must be taken in order to actually make use of the created models. The key steps here are plan deployment, plan monitoring and maintenance, the production of the final report, and review of the project [8].
Knowledge discovery uses data mining and machine learning techniques that have evolved through a synergy in artificial intelligence, computer science, statistics and other related fields [10].

3. Applications of data mining method in the energy system applications

Data mınıng method has been used by various researchers in energy system applications. This section deals with an overview of these applications. Some examples on the use of data mining method in the energy system applications are summarized in Table 2.

Area	Number of applications
Energy efficient building design	1
HVAC systems	2
Energy demand modeling	4
Electricity price forecast	1
Prediction of properties of refrigerants	3
Object tracking	2
Optimization of wind turbine	2
Cluster of load profiles	1
Modeling of absorption heat transformer	1
Analysis of fluidized-bed boiler	1

Table 2. Summary of numbers of applications presented in energy system applications

Küçüksille et al. have used data mining for the determination of thermophysical properties as the specific heat capacity, viscosity, heat conduction coefficient, density of the same refrigerants. This study presented ten modeling techniques within data mining process for the prediction of thermophysical properties of refrigerants (R134a, R404a, R407c and R410a). These are linear regression (LR), multi layer perception (MLP), pace regression (PR), simple linear regression (SLR), sequential minimal optimization (SMO), KStar, additive regression (AR), M5 model tree, decision table (DT), M5'Rules models. Relations depending on temperature and pressure were carried out for the determination of thermophysical properties as the specific heat capacity, viscosity, heat conduction coefficient, density of the refrigerants.
Obtained model results for every refrigerant were compared and the best model was investigated. Results indicate that use of derived formulations from these techniques will facilitate design and optimize of heat exchangers which is component of especially vapor compression refrigeration system [11].
Şencan has used data mining process to determine specific volume values of methanol/LiBr and methanol/LiCl used in absorption heat pump systems. Linear regression (LR), pace regression (PR), sequential minimal optimization (SMO), M5 model tree, M5'Rules and back propagation neural network (BPNN) models were applied within the data mining process. Mathematical formulations were found to be in good agreement with the experimental data [12].
Tso and Yau have used regression analysis, decision tree and neural networks models in the data mining approach for the prediction of electricity energy consumption. Model selection is based on the square root of average squared error. In an empirical application to an electricity energy consumption study, the decision tree and neural network models appear to be viable alternatives to the stepwise regression model in understanding energy consumption patterns and predicting energy consumption levels. With the emergence of the data mining approach for predictive modeling, different types of models can be built in a unified platform: to implement various modeling techniques, assess the performance of different models and select the most appropriate model for future prediction [13].
Kusiak et al. was applied data mining approach to analyze relationships among parameters of a circulating fluidized-bed boiler. The efficiency could be predicted to the same degree of accuracy with and without the data describing the fuel composition or boiler demand levels in study. Authors have determined that data mining approach is applicable to different types of burners and fuel types [14].

Figueiredo et al. have presented an electricity consumer characterization framework based on a knowledge discovery in databases procedure, supported by data mining techniques. This framework consists of two main modules: the load profiling module and the classification module. The load profiling module creates a set of consumer classes using a clustering operation and the representative load profiles for each class. The classification module uses this knowledge to build a classification model able to assign different consumers to the existing classes [15].

The electricity price forecast framework, which can predict the normal price as well as the price spikes based on data mining approach by Lu et al. has carried out. The proposed model is based on a mining database including market clearing price, trading hour, electricity demand, electricity supply and reserve. This proposed model is able to generate forecasted price spike, level of spike and associated forecast confidence level [16].

Şencan et al. used different methods such as linear regression (LR), pace regression (PR), sequential minimal optimization (SMO), M5 model tree, M50 rules, decision table and back propagation neural network (BPNN) for modelling the absorption heat transformer. A theoretical modeling of an absorption heat transformer for the temperature range obtained from an experimental solar pond with dimensions 3.5 x 3.5 x 2 m is presented. The working fluid pair in the absorption heat transformer is aqueous ternary hydroxide fluid consisting of sodium, potassium and caesium hydroxides in the proportions 40:36:24 (NaOH:KOH:CsOH). The best results were obtained by the back propagation neural network model. A new formulation based on the BPNN is presented to determine the flow ratio (FR) and the coefficient of performance (COP) of the absorption heat transformer [17].

Küçüksille et al. has applied data mining approach for the modeling of thermodynamic properties of alternative refrigerants. In addition, mathematical equations in order to calculate enthalpy, entropy and specific volume values of each refrigerant were presented.The values calculated from obtained formulations were found to be in good agreement with actual values. The results of this work show that DM can use for predicting accuracy of thermodynamic properties of refrigerants for every temperature and pressure [18].

Yu et al. used a decision tree method for building energy demand modeling. This method is applied to Japanese residential buildings for predicting and classifying building EUI levels and its basic steps, such as the generation of decision tree based on training data and the evaluation of decision tree based on test data are presented. The results have demonstrated that the use of decision tree method can classify and predict building energy demand levels accurately (93% for training data and 92% for test data), identify and rank significant factors of building EUI levels automatically, and provide the combination of significant factors as well as the threshold values that will lead to high building energy performance [19].

Kim et al. developed a process which can help project teams discover useful patterns to improve energy efficient building design. This paper utilized data mining technology, which is a data analysis process that combines different techniques from machine learning, pattern recognition, statistics, and visualization, to automatically extract concepts, interrelationships and patterns of interest from a large dataset. By applying data mining technology to the analysis of energy efficient building designs one can identify valid, useful, and previously unknown patterns out of energy simulation modeling [20].

Hou et al. used data mining (DM) method is developed to detect and diagnose sensor faults based on the past running performance data in heating, ventilating and air conditioning (HVAC) systems, combining a rough set approach and an artificial neural network (ANN). The reduced information is used to develop classification rules and train the neural network to infer appropriate parameters. The differences between measured thermodynamic states and predicted states obtained from models for normal performance (residuals) are used as performance indices for sensor fault detection and diagnosis. Real test results from a real HVAC system show that only the temperature and humidity measurements of many air handling units (AHU) can work very well as the measurements to distinguish simultaneous temperature sensor faults of the supply chilled water (SCW) and return chilled water (RCW). The logic diagram of the DM based sensor fault detection and diagnosis strategy is shown in Fig. 2 [21].

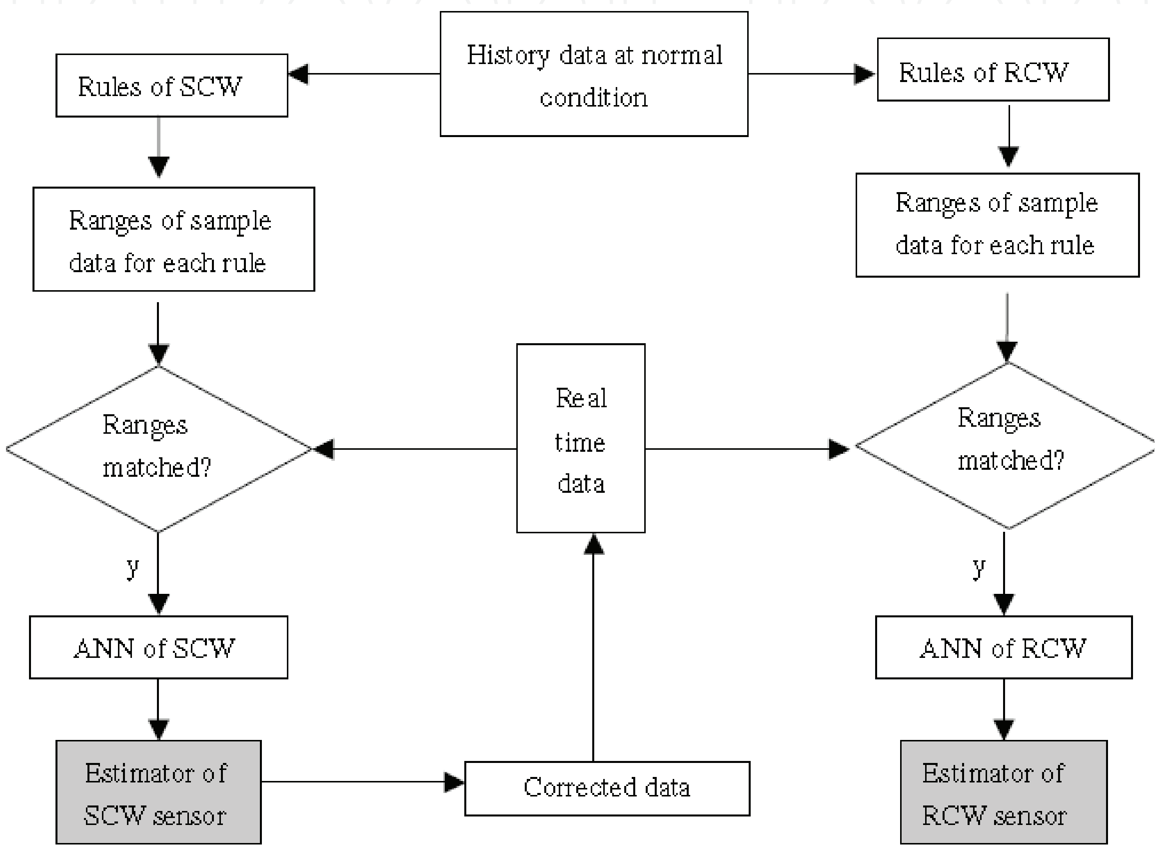

Fig. 2. Logic diagram of DM based sensor FDD strategy [21].

Dominguez-Navarro et al. used data mining to analyze the composition of the electric demand among the different consumption and the behavior of each type of load. The proposed method uses a heuristic optimization algorithm (Tabu Search) for minimizing the error between the real demand and the calculated approximation to this demand. This search is adaptative because the algorithm changes the relative weight of each load as well as the profile of each load. The obtained results show the good operation of the proposedmethodology [22].

Lu et al. used data mining based electricity price forecast framework, which can predict the normal price as well as the price spikes. The normal price can be predicted by a previously proposed wavelet and neural network based forecast model, while the spikes are forecasted based on a data mining approach. This paper focuses on the spike prediction and explores the reasons for price spikes based on the measurement of a proposed composite supply–demand balance index (SDI) and relative demand index (RDI). These indices are able to reflect the relationship among electricity demand, electricity supply and electricity reserve capacity. The proposed model is based on a mining database including market clearing price, trading hour, electricity demand, electricity supply and reserve. Bayesian classification and similarity searching techniques are used to mine the database to find out the internal relationships between electricity price spikes and these proposed. The mining results are used to form the price spike forecast model. This proposed model is able to generate forecasted price spike, level of spike and associated forecast confidence level. The model is tested with the Queensland electricity market data with promising results. Flow chart of the comprehensive electricity price forecast model in Fig. 3 were presented [23].

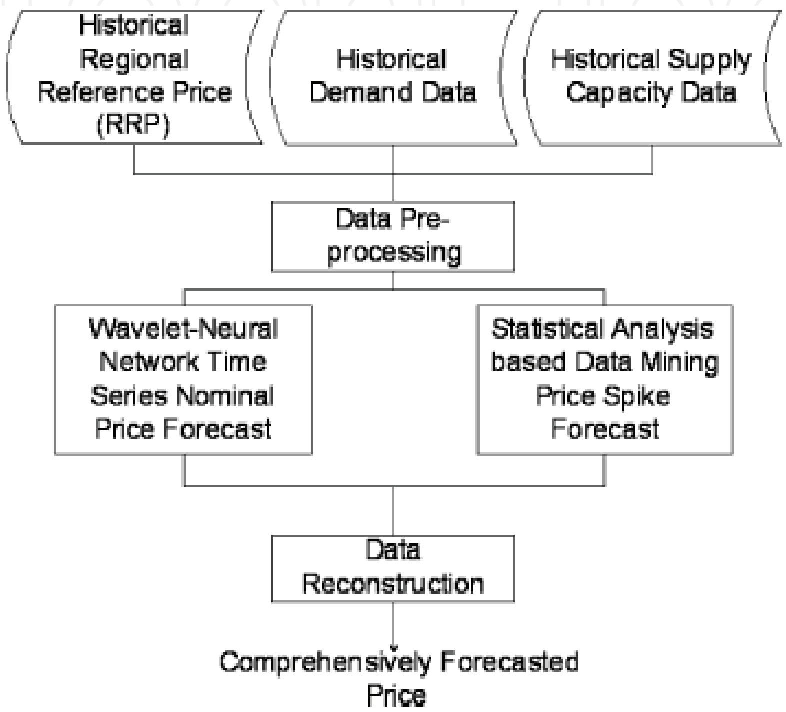

Fig. 3. Flow chart of the comprehensive electricity price forecast model [23].

Tseng and Lin proposed a novel data mining algorithm named TMP-Mine with a special data structure named TMP-Tree for efficiently discovering the temporal movement patterns of objects in sensor networks. They proposed novel location prediction strategies that utilize the discovered temporal movement patterns so as to reduce the prediction errors for energy savings. Through empirical evaluation on various simulation conditions and real dataset, TMP-Mine and the proposed prediction strategies are shown to deliver excellent performance in terms of scalability, accuracy and energy efficiency [24].

Tseng and Lu proposed a novel strategy named multi-level object tracking strategy (MLOT) for energy-efficient and real-time tracking of the moving objects in sensor networks by mining the movement log. In MLOT, they first conducted hierarchical clustering to form a hierarchical model of the sensor nodes. Second, the movement logs of the moving objects are analyzed by a data mining algorithm to obtain the movement patterns, which are then used to predict the next position of a moving object. They used the multi-level structure to represent the hierarchical relations among sensor nodes so as to achieve the goal of keeping track of moving objects in a real-time manner. Through experimental evaluation of various simulated conditions, the proposed method is shown to deliver excellent performance in terms of both energy efficiency and timeliness. Fig. 4 shows the system architecture. The

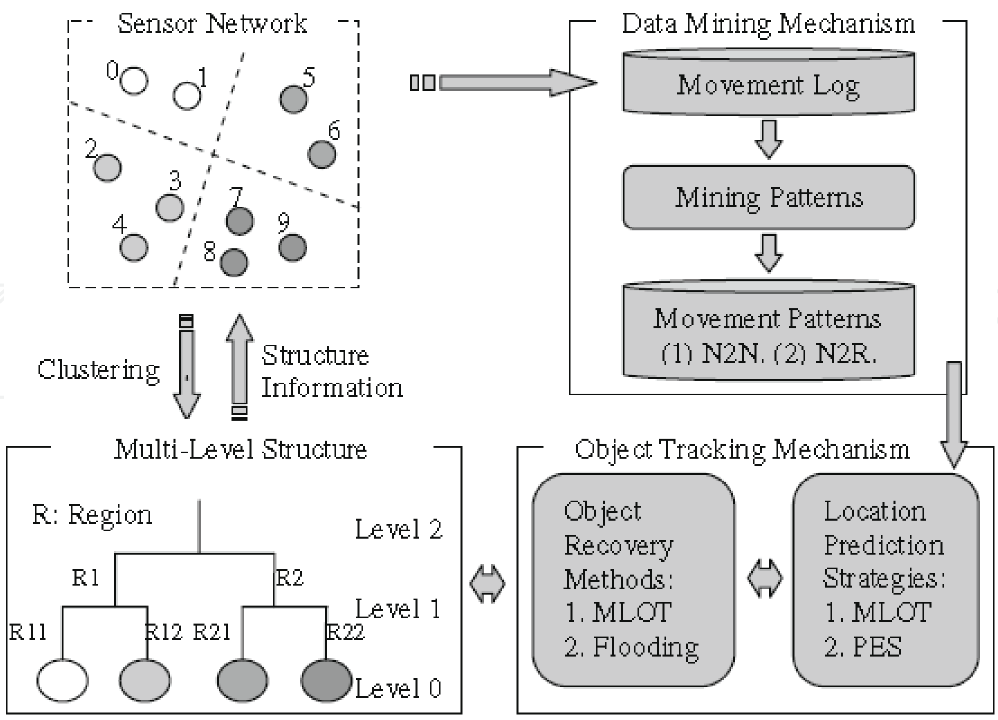

Fig. 4. System architecture [25].

system workflow consists of three main phases: (1) clustering of sensor nodes; (2) discovery of movement patterns; and (3) prediction and recovery of locations of moving objects [25].

Azadeh et al. presented an integrated fuzzy system, data mining and time series framework to estimate and predict electricity demand for seasonal and monthly changes in electricity consumption especially in developing countries such as China and Iran with non-stationary data. It is difficult to model uncertain behavior of energy consumption with only conventional fuzzy system or time series and the integrated algorithm could be an ideal substitute for such cases. To construct fuzzy systems, a rule base is needed. Because a rule base is not available, for the case of demand function, look up table which is one of the extracting rule methods is used to extract the rule base. This system is defined as FLT. Also, decision tree method which is a data mining approach is similarly utilized to extract the rule base. This system is defined as FDM. Preferred time series model is selected from linear (ARMA) and nonlinear model. For this, after selecting preferred ARMA model, McLeod–Li test is applied to determine nonlinearity condition. When, nonlinearity condition is satisfied, preferred nonlinear model is selected and compare with preferred ARMA model and finally one of this is selected as time series model. At last, ANOVA is used for selecting preferred model from fuzzy models and time series model. Also, the impact of data preprocessing and postprocessing on the fuzzy system performance is considered by the algorithm. In addition, another unique feature of the proposed algorithm is utilization of autocorrelation function (ACF) to define input variables, whereas conventional methods which use trial and error method. Monthly electricity consumption of Iran from 1995 to 2005 is considered as the case of this study. The MAPE estimation of genetic algorithm (GA), artificial neural network (ANN) versus the proposed algorithm shows the appropriateness of the proposed algorithm [26].

Kusiak et al. presented data-driven approach for minimization of the energy to air condition a typical office-type facility. Eight data-mining algorithms are applied to model the nonlinear relationship among energy consumption, control settings (supply air temperature and supply air static pressure), and a set of uncontrollable parameters. The multiple-linear perceptron (MLP) ensemble outperforms other models tested in this research, and therefore it is selected to model a chiller, a pump, a fan, and a reheat device. These four models are integrated into an energy optimization model with two decision variables, the setpoint of the supply air temperature and the static pressure in the air handling unit. The model is solved with a particle swarm optimization algorithm. The optimization results have demonstrated the total energy consumed by the heating, ventilation, and air-conditioning system is reduced by over 7%. Fig. 7 shows the flowchart of the Particle swarm optimization (PSO) algorithm [27].

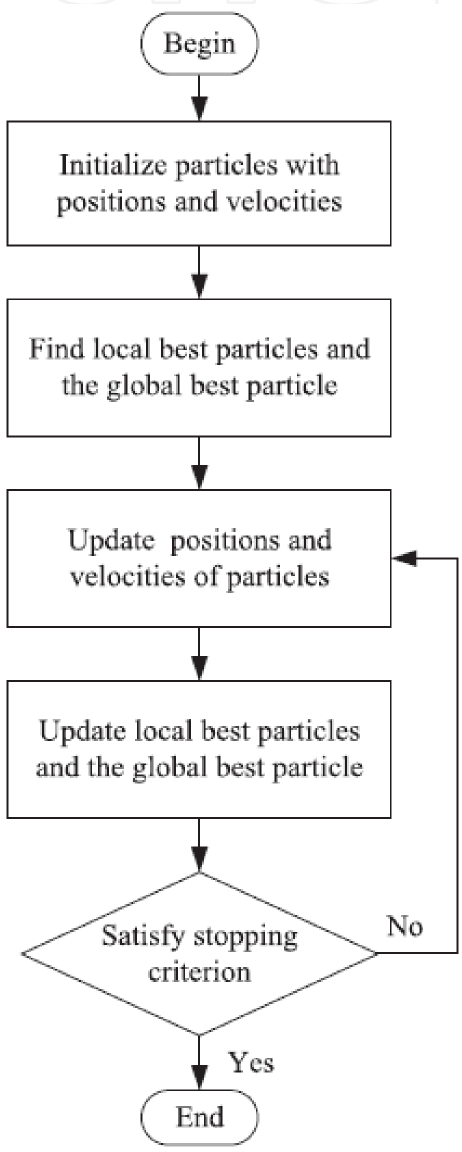

Fig. 7. Flowchart of the PSO algorithm [27].

Kusiak and Zheng presented an evolutionary computation approach for optimization of power factor and power output of wind turbines. Data-mining algorithms capture the relationships among the power output, power factor, and controllable and non-controllable variables of a 1.5 MW wind turbine. An evolutionary strategy algorithm solves the data-derived optimization model and determines optimal control settings. Computational experience has demonstrated opportunities to improve the power factor and the power output by optimizing set points of blade pitch angle and generator torque. It is shown that the pitch angle and the generator torque can be controlled to maximize the energy capture from the wind and enhance the quality of the power produced by the wind turbine with a DFIG generator. These improvements are in the presence of reactive power remedies used in modern wind turbines. The concepts proposed in this paper are illustrated with the data collected at an industrial wind farm [28].

Kusiak et al. presented a data-driven approach for maximization of the power produced by wind turbines. The power optimization objective is accomplished by computing optimal control settings of wind turbines using data mining and evolutionary strategy algorithms. Data mining algorithms identify a functional mapping between the power output and controllable and non-controllable variables of a wind turbine. An evolutionary strategy algorithm is applied to determine control settings maximizing the power output of a turbine based on the identified model. Computational studies have demonstrated meaningful opportunities to improve the turbine power output by optimizing blade pitch and yaw angle. It is shown that the pitch angle is an important variable in maximizing energy captured from the wind. Power output can be increased by optimization of the pitch angle. The concepts proposed in this paper are illustrated with industrial wind farm data. Fig. 8 show optimization framework [29].

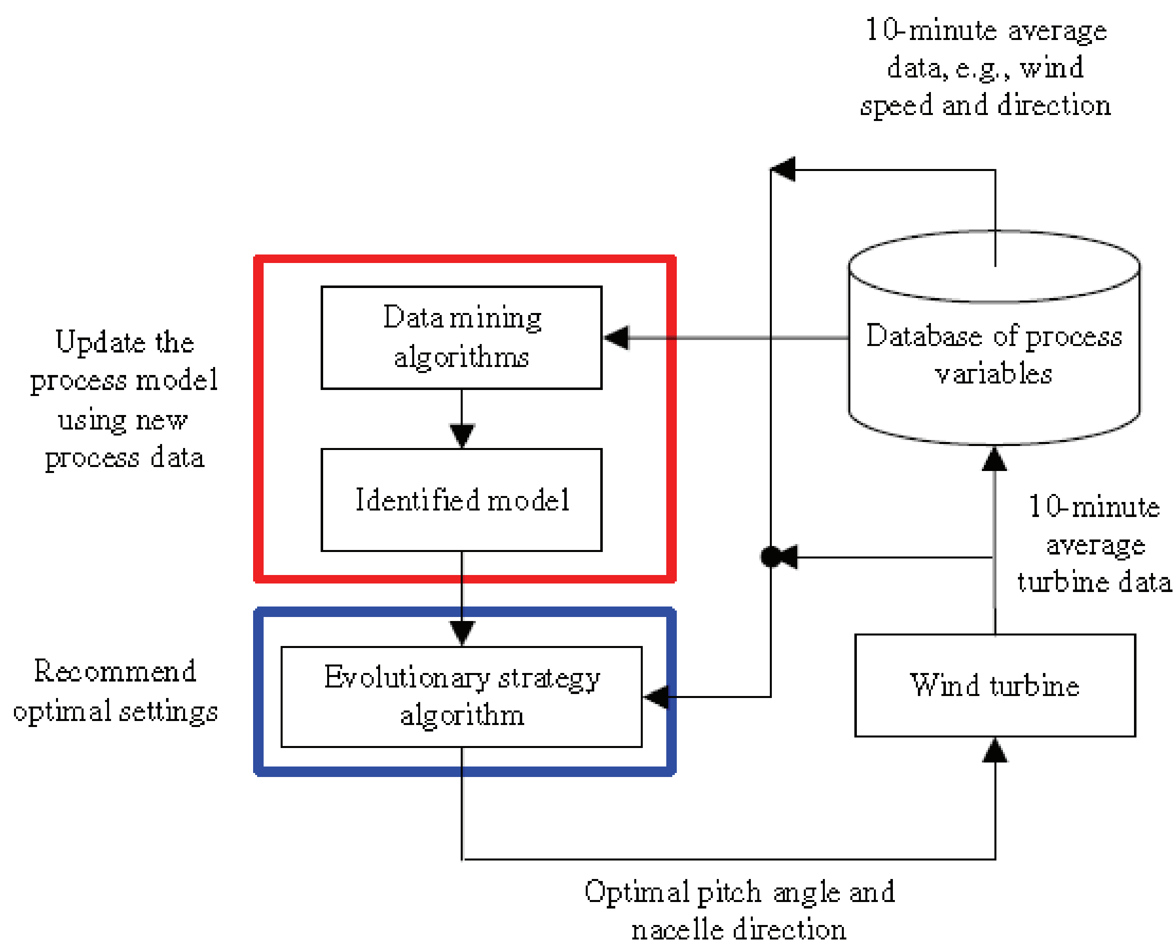

Fig. 8. Optimization framework [29].

Lin et al. applied datamining techniques to the CLP (China Light and Power) Power database in order to analyze (a) the effect of temperature and relative humidity on the peak load, (b) the cluster of load profiles in various buses in response to disturbances. The results draw attention to special phenomena associated with the particular system operation, such as that several substations whose load behavior during disturbances whenever they happened are statistically very similar to that of substation AAA are always very seriously impacted by the disturbances. Such results pinpoint these substations for further system studies, which may lead to enhanced overall performance. On the other hand, based on the limited records in the data-mining process, certain 'unexpected' findings are revealed (including in substation BBB) and closer scrutiny of future data collected in the associated buses will thus be called for [30].

4. Conclusions

From the description of the various applications presented in this paper, one can see that data mining method have been applied in many fields of energy systems. In this chapter, various applications made using data mining method have been reviewed. Available literature summaries published in this area is also presented. Data mining method is becoming useful as alternate approaches to conventional techniques. Data mining has also been applied for modeling, optimization, prediction and control of complex systems. As can be seen from the applications presented, data mining method has been applied successfully in a wide range of energy system applications.
Surely, the number of applications presented here is neither complete nor exhaustive but merely a sample of applications that demonstrate the usefulness and possible applications of data mining method. Based on the work presented here it is believed that data mining method offers an alternative method.

5. References

[1] Fayyad, U, Piatetsky-Shapiro, G and Smyth, P, 1996d, Knowledge discovery and data mining: towards a unifying framework. In Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, Portland, OR, pp. 82–88.

[2] Cabena, P, Hadjinian, P, Stadler, R, Verhees, J and Zanasi, A, 1998, Discovering Data Mining: From Concepts to Implementation. Prentice Hall.

[3] Hui, S. C., Jha, G., "Data mining for customer service support", Information & Management, Volume 38, Issue 1,2000, pp. 1-13

[4] Kurgan, L. A., Musilek, P., "A survey of Knowledge Discovery and Data Mining process models", The Knowledge Engineering Review, 2006,21:1, pp. 1 -24.

[5] Chapman et al., 2000 Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearar, C., et al. (2000). CRISP-DM 1.0 Step-by-Step Data Mining Guide (p. 13).

[6] Fernandez et al., 2002 I.B. Fernandez, S. H. Zanakis and S. Walczak, Knowledge discovery techniques for predicting country investment risk, Computers & Industrial Engineering (43) (2002), pp. 787–800.

[7] Wirth and Hipp, 2000 Wirth, R., & Hipp, J. (2000). CRIPS-DM: Towards a standard process model for data mining. In Proceedings of the fourth international conference on the practical applications of knowledge discovery and data mining, Manchester, UK (pp. 29–39).

[8] Shearar, 2000 C. Shearar, The CRISP-DM model: The new blueprint for data mining, Journal of Data Ware Housing 5 (4) (2000), p. 13-19.
[9] Tang and MacLennan, 2005 Z. Tang and J. MacLennan, Data mining with sql server 2005, Wiley (2005).
[10] Sencan, 2007 A. Sencan, Modeling of thermodynamic properties of refrigerant/absorbent couples using data mining process, Energy Conversion and Management 48 (2007), pp. 470–480.
[11] E. U. Küçüksille, R. Selbas, A. Şencan, Data mining techniques for thermophysical properties of refrigerants, Energy Conversion and Management 50 (2009) 399–412.
[12] A. Şencan, Modeling of thermodynamic properties of refrigerant/absorbent couples using Data Mining Process, Energy Conversion and Management, 48 (2007) 470–480.
[13] G. K. F. Tso, K. K. W. Yau, Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural Networks, Energy, 32 (2007) 1761–1768.
[14] A. Kusiak, A. Burns, F. Milster, Optimizing combustion efficiency of a circulating fluidized boiler: A data mining approach, International Journal of Knowledge Based Intelligent Engineering Systems, 9(2005) 263-274.
[15] V. Figueiredo, F. Rodrigues, J. G. Gouveia, An Electric Energy Consumer Characterization Framework Based on Data Mining Techniques, IEEE Transactions on Power Systems, 20(2005) 596-602.
[16] X. Lu, Z. Y. Dong, X. Li, Electricity market price spike forecast with data mining techniques, Electric Power Systems Research, 73 (2005) 19–29.
[17] A. Sencan, O.Kızılkan, N.C. Bezir, S.A. Kalogirou, Different methods for modeling absorption heat transformer powered by solar pond, Energy Conversion and Management 48 (2007) 724–735.
[18] E. U. Küçüksille, R. Selbas, A. Şencan, Prediction of thermodynamic properties of refrigerants using data mining, Energy Conversion and Management, In Press.
[19] Zhun Yu, Fariborz Haghighat, Benjamin C.M. Fung, Hiroshi Yoshino, Energy and Buildings 42 (2010) 1637–1646.
[20] Hyunjoo Kim, Annette Stumpf , Wooyoung Kim, Analysis of an energy efficient building design through data mining approach, Automation in Construction, In Press.
[21] Zhijian Hou, Zhiwei Lian, Ye Yao, Xinjian Yuan, Data mining based sensor fault diagnosis and validation for building air conditioning system,Energy Conversion and Management 47 (2006) 2479–2490.
[22] José A. Dominguez-Navarro, José L. Bernal-Agustín, Rodolfo Dufo-López, Data mining methodology for disaggregation of load demand, Electric Power Systems Research 79 (2009) 1393–1399.
[23] Xin Lu, Zhao Yang Dong, Xue Li, Electricity market price spike forecast with data mining techniques, Electric Power Systems Research 73 (2005) 19–29.
[24] Vincent S. Tseng, Kawuu W. Lin, Energy efficient strategies for object tracking in sensor networks: A data mining approach, The Journal of Systems and Software 80 (2007) 1678–1698.

[25] Vincent S. Tseng, Eric Hsueh-Chan Lu, Energy-efficient real-time object tracking in multi-level sensor networks by mining and predicting movement patterns, The Journal of Systems and Software 82 (2009) 697–706.
[26] A. Azadeh, M. Saberi, S. F. Ghaderi, A. Gitiforouz , V. Ebrahimipour, Improved estimation of electricity demand function by integration of fuzzy system and data mining approach, Energy Conversion and Management 49 (2008) 2165–2177.
[27] Andrew Kusiak, Mingyang Li, Fan Tang, Modeling and optimization of HVAC energy consumption, Applied Energy 87 (2010) 3092–3102.
[28] Andrew Kusiak, Haiyang Zheng, Optimization of wind turbine energy and power factor with an evolutionary computation algorithm, Energy 35 (2010) 1324–1332.
[29] Andrew Kusiak, Haiyang Zheng, Zhe Song, Power optimization of wind turbines with data mining and evolutionary computation, Renewable Energy 35 (2010) 695–702.
[30] J. K. Lin, S. K. Tso, H. K. Ho, C. M. Mak, K. M. Yung, Y. K. Ho, Study of climatic effects on peak load and regional similarity of load profiles following disturbances based on data mining, Electrical Power and Energy Systems 28 (2006) 177–185.

Knowledge-Oriented Applications in Data Mining

Edited by Prof. Kimito Funatsu

ISBN 978-953-307-154-1

Hard cover, 442 pages

Publisher InTech

Published online 21, January, 2011

Published in print edition January, 2011

The progress of data mining technology and large public popularity establish a need for a comprehensive text on the subject. The series of books entitled by 'Data Mining' address the need by presenting in-depth description of novel mining algorithms and many useful applications. In addition to understanding each section deeply, the two books present useful hints and strategies to solving problems in the following chapters. The contributing authors have highlighted many future research directions that will foster multi-disciplinary collaborations and hence will lead to significant development in the field of data mining.