American Journal of Mobile Systems, Applications and Services
Vol. 1, No. 3, 2015, pp. 208-217
http://www.aiscience.org/journal/ajmsas

Proposing a Criterion for Clusters Validation by Combined Hidden Markov Model and Clustering Method

Sergio Komini*, Margaret Siredon

School of Computer Science, Universidad de Chile, Santiago, Chile

Abstract

In this paper a new criterion for clusters validation is proposed. This new cluster validation criterion is used to approximate the goodness of a cluster. A clustering ensmble framework based on the new metric is proposed. The main idea behind the framework is to extract the most stable clusters in terms of the defined criteria. The term is a misnomer, because the goal is the extraction of patterns and knowledge from large amount of data, not the extraction of data itself. It also is a buzzword and is also frequently applied to any form of large-scale data or information processing (collection, extraction, warehousing, analysis, and statistics) as well as any application of computer decision support system, including artificial intelligence, machine learning, and business intelligence. In this paper we proposed a combined model of HMM and C-Mediod to Item dataset clustering.

Keywords

Clustering Ensemble, Stability Measure, Extended EAC, Hidden Markov Model, Sequences, C-Mediod, Data Mining.

Received: August 20, 2015 / Accepted: December 2, 2015 / Published online: January 11, 2016

1. Introduction

A hidden Markov model (HMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states. An HMM can be presented as the simplest dynamic Bayesian network. The mathematics behind the HMM was developed by L. E. Baum and coworkers. [1-5] It is closely related to an earlier work on the optimal nonlinear filtering problem by Ruslan L. Stratonovich [6] who was the first to describe the forward-backward procedure.

In simpler Markov models (like a Markov chain), the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a hidden Markov model, the state is not directly visible, but output, dependent on the state, is visible. Each state has a probability distribution over the possible output tokens. Therefore the sequence of tokens generated by an HMM gives some information about the sequence of states. Note that the adjective 'hidden' refers to the state sequence through which the model passes, not to the parameters of the model; the model is still referred to as a 'hidden' Markov model even if these parameters are known exactly.

Hidden Markov models are especially known for their application in temporal pattern recognition such as speech, handwriting, gesture recognition [7-10] part-of-speech tagging, musical score following [11] partial discharges [12] and bioinformatics.

In its discrete form, a hidden Markov process can be visualized as a generalization of the Urn problem with replacement (where each item from the urn is returned to the original urn before the next step) [13, 14, 15], [16-20] Consider this example: in a room that is not visible to an observer there is a genie. The room contains urns X1, X2,

* Corresponding author
E-mail address: sergio.k34@hotmail.com (S. Komini)

X3, ... each of which contains a known mix of balls, each ball labeled y1, y2, y3, ... The genie chooses an urn in that room and randomly draws a ball from that urn. It then puts the ball onto a conveyor belt, where the observer can observe the sequence of the balls but not the sequence of urns from which they were drawn. The genie has some procedure to choose urns; the choice of the urn for the n-th ball depends only upon a random number and the choice of the urn for the (n − 1)-th ball. The choice of urn does not directly depend on the urns chosen before this single previous urn; therefore, this is called a Markov process.

The Markov process itself cannot be observed, and only the sequence of labeled balls can be observed, thus this arrangement is called a "hidden Markov process". This is illustrated by the lower part of the diagram shown in Figure 1, where one can see that balls y1, y2, y3, y4 can be drawn at each state. Even if the observer knows the composition of the urns and has just observed a sequence of three balls, e.g. y1, y2 and y3 on the conveyor belt, the observer still cannot be sure which urn (i.e., at which state) the genie has drawn the third ball from. However, the observer can work out other information, such as the likelihood that the third ball came from each of the urns [21, 22].

A hidden Markov model can be considered a generalization of a mixture model where the hidden variables (or latent variables), which control the mixture component to be selected for each observation, are related through a Markov process [23-27, 130-140] rather than independent of each other. Recently, hidden Markov models have been generalized to pairwise [28, 29] Markov models and triplet Markov models which allow consideration of more complex data structures [30-38] and the modelling of nonstationary data [39] in large data sets. It bridges the gap from applied statistics and artificial intelligence (which usually provide the mathematical background) to database management by exploiting the way data is stored and indexed in databases to execute the actual learning and discovery algorithms more efficiently, allowing such methods to be applied to ever larger data sets.

The term is a misnomer, because the goal is the extraction of patterns and knowledge from large amount of data, not the extraction of data itself [40, 41], [42-45] It also is a buzzword [46-50] and is also frequently applied to any form of large-scale data or information processing (collection, extraction, warehousing, analysis, and statistics) as well as any application of computer decision support system, including artificial intelligence, machine learning, and business intelligence. The popular book "Data mining: Practical machine learning tools and techniques with Java [51] (which covers mostly machine learning material) was originally to be named just "Practical machine learning", and the term "data mining" was only added for marketing reasons [52, 53, 54] Often the more general terms "(large scale) data analysis", or "analytics" – or when referring to actual methods, artificial intelligence and machine learning-are more appropriate [55-59, 120-137].

The actual data mining task is the automatic or semi-automatic analysis of large quantities of data to extract previously unknown interesting patterns such as groups of data records (cluster analysis), unusual records (anomaly detection) and dependencies (association rule mining). This usually involves using database techniques such as spatial indices. These patterns can then be seen as a kind of summary of the input data, and may be used in further analysis or, for example, in machine learning and predictive analytics. For example, the data mining step might identify multiple groups in the data, which can then be used to obtain more accurate prediction results by a decision support system. Neither the data collection, data preparation, nor result interpretation and reporting are part of the data mining step, but do belong to the overall KDD process as additional steps [60, 61].

The related terms data dredging, data fishing, and data snooping refer to the use of data mining methods to sample parts of a larger population data set that are (or may be) too small for reliable statistical inferences to be made about the validity of any patterns discovered. These methods can, however, be used in creating new hypotheses to test against the larger data populations.

Research and evolution

The premier professional body in the field is the ssociation for Computing Machinery's (ACM) Special Interest Group (SIG) on Knowledge Discovery and Data Mining (SIGKDD) [11, 12] Since 1989 this ACM SIG has hosted an annual international conference and published its proceedings [62-66] and since 1999 it has published a biannual academic journal titled “SIGKDD Explorations” [67-72].

2. Model

Each oval shape represents a random variable that can adopt any of a number of values. The random variable x(t) is the hidden state at time t (with the model from the above diagram, x(t) ∋ {x1, x2, x3}). The random variable y(t) is the observation at time t (with y(t) ∋ {y1, y2, y3, y4}). The arrows in the diagram (often called a trellis diagram) denote conditional dependencies.

From the diagram, it is clear that the conditional probability distribution of the hidden variable x(t) at time t, given the values of the hidden variable x at all times, depends only on the value of the hidden variable x(t − 1): the values at time t − 2 and before have no influence. This is called the Markov property. Similarly, the value of the observed variable y(t) only depends on the value of the hidden variable x(t) (both at time.

In the standard type of hidden Markov model considered here, the state space of the hidden variables is discrete, while the observations themselves can either be discrete (typically generated from a categorical distribution) or continuous (typically from a Gaussian distribution). The parameters of a hidden Markov model are of two types, transition probabilities and emission probabilities (also known as output probabilities). The transition probabilities control the way the hidden state at time t is chosen given the hidden state at time t-1.

The hidden state space is assumed to consist of one of N possible values, modeled as a categorical distribution. (See the section below on extensions for other possibilities.) This means that for each of the N possible states that a hidden variable at time t can be in, there is a transition probability from this state to each of the N possible states of the hidden variable at time t+1, for a total of N^2 transition probabilities. Note that the set of transition probabilities for transitions from any given state must sum to 1. Thus, the $N \times N$ matrix of transition probabilities is a Markov matrix.

In addition, for each of the N possible states, there is a set of emission probabilities governing the distribution of the observed variable at a particular time given the state of the hidden variable at that time. The size of this set depends on the nature of the observed variable. For example, if the observed variable is discrete with M possible values, governed by a categorical distribution, there will be M-1 separate parameters, for a total of N(M-1) emission parameters over all hidden states. On the other hand, if the observed variable is an M-dimensional vector distributed according to an arbitrary multivariate Gaussian distribution, there will be M parameters controlling the means and M(M+1)/2 parameters controlling the covariance matrix, for a total of $N(M + \frac{M(M+1)}{2}) = NM(M+3)/2 = O(NM^2)$ emission parameters. (In such a case, unless the value of M is small, it may be more practical to restrict the nature of the covariances between individual elements of the observation vector, e.g. by assuming that the elements are independent of each other, or less restrictively, are independent of all but a fixed number of adjacent elements.)

Polls conducted in 2002, 2004, and 2007 show that the CRISP-DM methodology is the leading methodology used by data miners [73-81] The only other data mining standard named in these polls was SEMMA. However, 3-4 times as many people reported using CRISP-DM. Several teams of researchers have published reviews of data mining process models and Azevedo and Santos conducted a comparison of CRISP-DM and SEMMA in 2008 [82].

Pre-processing

Before data mining algorithms can be used, a target data set must be assembled. As data mining can only uncover patterns actually present in the data, the target data set must be large enough to contain these patterns while remaining concise enough to be mined within an acceptable time limit. A common source for data is a data mart or data warehouse. Pre-processing is essential to analyze the multivariate data sets before data mining. The target set is then cleaned. Data cleaning removes the observations containing noise and those with missing data.

Data mining

Data mining involves six common classes of tasks [83]:

Anomaly detection (Outlier/change/deviation detection) – The identification of unusual data records, that might be interesting or data errors that require further investigation.

Association rule learning (Dependency modelling) – Searches for relationships between variables. For example a supermarket might gather data on customer purchasing habits. Using association rule learning, the supermarket can determine which products are frequently bought together and use this information for marketing purposes. This is sometimes referred to as market basket analysis.

Clustering–is the task of discovering groups and structures in the data that are in some way or another "similar", without using known structures in the data.

Classification–is the task of generalizing known structure to apply to new data. For example, an e-mail program might attempt to classify an e-mail as "legitimate" or as "spam."

Regression–attempts to find a function which models the data with the least error.

Summarization–providing a more compact representation of the data set, including visualization and report generation.

Results validation

Data mining can unintentionally be misused, and can then produce results which appear to be significant; but which do not actually predict future behavior and cannot be reproduced on a new sample of data and bear little use. Often this results from investigating too many hypotheses and not performing proper statistical hypothesis testing. A simple version of this problem in machine learning is known as overfitting, but the same problem can arise at different phases of the process and thus a train/test split - when applicable at all - may not be sufficient to prevent this from happening.

3. Clustering Methods

Clustering algorithms can be categorized based on their cluster model, as listed above. The following overview will only list the most prominent examples of clustering algorithms, as there are possibly over 100 published clustering algorithms. Not all provide models for their clusters and can thus not easily be categorized. An overview of algorithms explained in Wikipedia can be found in the list of statistics algorithms.

There is no objectively "correct" clustering algorithm, but as it was noted, "clustering is in the eye of the beholder [84]. The most appropriate clustering algorithm for a particular problem often needs to be chosen experimentally, unless there is a mathematical reason to prefer one cluster model over another. It should be noted that an algorithm that is designed for one kind of model has no chance on a data set that contains a radically different kind of model. For example, k-means cannot find non-convex clusters.

20 clusters extracted, most of which contain single elements, since linkage clustering does not have a notion of "noise". Many people used this method on the different.

K-means has a number of interesting theoretical properties. On the one hand, it partitions the data space into a structure known as a Voronoi diagram. On the other hand, it is conceptually close to nearest neighbor classification, and as such is popular in machine learning.

Distribution-based clustering produces complex models for clusters that can capture correlation and dependence between attributes. However, these algorithms put an extra burden on the user: for many real data sets, there may be no concisely defined mathematical model (figure 1). Many people used this method on the different application [85, 86-100].

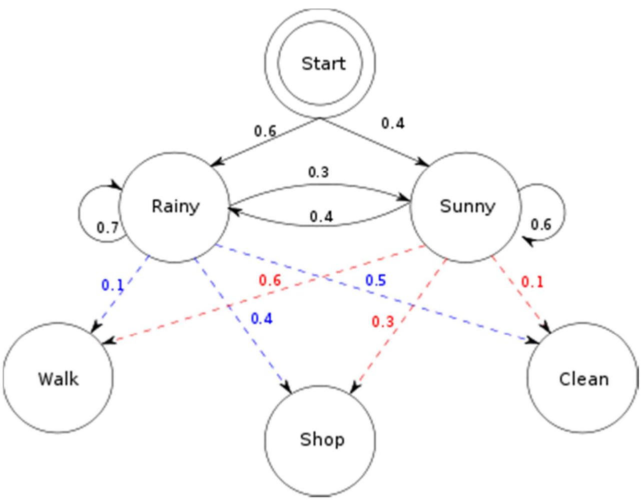

Figure 1. HMM for association mining.

4. Using Application

Business

In business, data mining is the analysis of historical business activities, stored as static data in data warehouse databases. The goal is to reveal hidden patterns and trends. Data mining software uses advanced pattern recognition algorithms to sift through large amounts of data to assist in discovering previously unknown strategic business information. Examples of what businesses use data mining for include performing market analysis to identify new product bundles, finding the root cause of manufacturing problems, to prevent customer attrition and acquire new customers, cross-sell to existing customers, and profile customers with more accuracy.

Science and engineering

In recent years, data mining has been used widely in the areas of science and engineering, such as bioinformatics, genetics, medicine, education and electrical power engineering.

In the study of human genetics, sequence mining helps address the important goal of understanding the mapping relationship between the inter-individual variations in human DNA sequence and the variability in disease susceptibility. In simple terms, it aims to find out how the changes in an individual's DNA sequence affects the risks of developing common diseases such as cancer, which is of great importance to improving methods of diagnosing, preventing, and treating these diseases. One data mining method that is used to perform this task is known as multifactor dimensionality reduction.

Medical data mining

In 2011, the case of Sorrell v. IMS Health, Inc., decided by the Supreme Court of the United States, ruled that pharmacies may share information with outside companies. This practice was authorized under the 1st Amendment of the Constitution, protecting the "freedom of speech. However, the passage of the Health Information Technology for Economic and Clinical Health Act (HITECH Act) helped to initiate the adoption of the electronic health record (EHR) and supporting technology in the United States. The HITECH Act was signed into law on February 17, 2009 as part of the American Recovery and Reinvestment Act (ARRA) and helped to open the door to medical data mining. Prior to the signing of this law, estimates of only 20% of United States based physician were utilizing electronic patient records. Søren Brunak notes that "the patient record becomes as information-rich as possible" and thereby "maximizes the data mining opportunities." Hence, electronic patient records further expands the possibilities regarding medical data mining thereby opening the door to a vast source of medical data analysis.

Surveillance

Data mining has been used by the U.S. government. Programs include the Total Information Awareness (TIA) program, Secure Flight (formerly known as Computer-Assisted Passenger Prescreening System (CAPPS II)), Analysis, Dissemination, Visualization, Insight, Semantic Enhancement (ADVISE), and the Multi-state Anti-Terrorism Information Exchange (MATRIX). These programs have been discontinued due to controversy over whether they violate the 4th Amendment to the United States Constitution, although many programs that were formed under them continue to be funded by different organizations or under different names.

In the context of combating terrorism, two particularly plausible methods of data mining are "pattern mining" and "subject-based data mining."

Pattern mining

"Pattern mining" is a data mining method that involves finding existing patterns in data. In this context patterns often means association rules. The original motivation for searching association rules came from the desire to analyze supermarket transaction data, that is, to examine customer behavior in terms of the purchased products. For example, an association rule "beer ⇒ potato chips (80%)" states that four out of five customers that bought beer also bought potato chips.

In the context of pattern mining as a tool to identify terrorist activity, the National Research Council provides the following definition: "Pattern-based data mining looks for patterns (including anomalous data patterns) that might be associated with terrorist activity — these patterns might be regarded as small signals in a large ocean of noise." Pattern Mining includes new areas such a Music Information Retrieval (MIR) where patterns seen both in the temporal and non temporal domains are imported to classical knowledge discovery search methods.

Subject-based data mining

"Subject-based data mining" is a data mining method involving the search for associations between individuals in data. In the context of combating terrorism, the National Research Council provides the following definition: "Subject-based data mining uses an initiating individual or other datum that is considered, based on other information, to be of high interest, and the goal is to determine what other persons or financial transactions or movements, etc., are related to that initiating datum.

5. Extension of Model

In the hidden Markov models considered above, the state space of the hidden variables is discrete, while the observations themselves can either be discrete (typically generated from a categorical distribution) or continuous (typically from a Gaussian distribution). Hidden Markov models can also be generalized to allow continuous state spaces. Examples of such models are those where the Markov process over hidden variables is a linear dynamical system, with a linear relationship among related variables and where all hidden and observed variables follow a Gaussian distribution. In simple cases, such as the linear dynamical system just mentioned, exact inference is tractable (in this case, using the Kalman filter); however, in general, exact inference in HMMs with continuous latent variables is infeasible, and approximate methods must be used, such as the extended Kalman filter or the particle filter.

Hidden Markov models are generative models, in which the joint distribution of observations and hidden states, or equivalently both the prior distribution of hidden states (the transition probabilities) and conditional distribution of observations given states (the emission probabilities), is modeled. The above algorithms implicitly assume a uniform prior distribution over the transition probabilities. However, it is also possible to create hidden Markov models with other types of prior distributions. An obvious candidate, given the categorical distribution of the transition probabilities, is the Dirichlet distribution, which is the conjugate prior distribution of the categorical distribution. Typically, a symmetric Dirichlet distribution is chosen, reflecting ignorance about which states are inherently more likely than others. The single parameter of this distribution (termed the concentration parameter) controls the relative density or sparseness of the resulting transition matrix. A choice of 1 yields a uniform distribution. Values greater than 1 produce a dense matrix, in which the transition probabilities between pairs of states are likely to be nearly equal. Values less than 1 result in a sparse matrix in which, for each given source state, only a small number of destination states have non-negligible transition probabilities. It is also possible to use a two-level prior Dirichlet distribution, in which one Dirichlet distribution (the upper distribution) governs the parameters of another Dirichlet distribution (the lower distribution), which in turn governs the transition probabilities. The upper distribution governs the overall distribution of states, determining how likely each state is to occur; its concentration parameter determines the density or sparseness of states. Such a two-level prior distribution, where both concentration parameters are set to produce sparse distributions, might be useful for example in unsupervised part-of-speech tagging, where some parts of speech occur much more commonly than others; learning algorithms that assume a uniform prior distribution generally perform poorly on this task. The parameters of models of this sort, with non-uniform prior distributions, can be learned using Gibbs sampling or extended versions of the expectation-maximization algorithm.

Another recent extension is the triplet Markov model, in which an auxiliary underlying process is added to model some data specificities. Many variants of this model have been proposed. One should also mention the interesting link that has been established between the theory of evidence and the triplet Markov models and which allows to fuse data in Markovian context and to model nonstationary data.

While the term "data mining" itself has no ethical implications, it is often associated with the mining of information in relation to peoples' behavior (ethical and otherwise) [86, 100-137].

The ways in which data mining can be used can in some cases and contexts raise questions regarding privacy, legality, and ethics. In particular, data mining government or commercial data sets for national security or law enforcement purposes, such as in the Total Information Awareness Program or in ADVISE, has raised privacy.

Data mining requires data preparation which can uncover information or patterns which may compromise confidentiality and privacy obligations. A common way for this to occur is through data aggregation. Data aggregation involves combining data together (possibly from various sources) in a way that facilitates analysis (but that also might make identification of private, individual-level data deducible or otherwise apparent). This is not data mining per se, but a result of the preparation of data before – and for the purposes of – the analysis. The threat to an individual's privacy comes into play when the data, once compiled, cause the data miner, or anyone who has access to the newly compiled data set, to be able to identify specific individuals, especially when the data were originally anonymous [87, 88, 89].

6. Conclusion

Generally, data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. In simpler Markov models (like a Markov chain), the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a hidden Markov model, the state is not directly visible, but output, dependent on the state, is visible. Each state has a probability distribution over the possible output tokens. Therefore the sequence of tokens generated by an HMM gives some information about the sequence of states. Note that the adjective 'hidden' refers to the state sequence through which the model passes, not to the parameters of the model; the model is still referred to as a 'hidden' Markov model even if these parameters are known exactly. We proposed a fuzzy HMM for Data Mining, which applied on several applications. Our results showed that our proposed method is better than current methods.

References

[1] K. G. Parthiban and S. Vijayachitra, Spike Detection from Electroencephalogram Signals with Aid of Hybrid Genetic Algorithm-Particle Swarm Optimization, Journal of Medical Imaging and Health Informatics, 5, 936-944 (2015).

[2] I. Barwal and S. C. Yadav, Rebaudioside A Loaded Poly-D, L-Lactide-Nanoparticles as an Anti-Diabetic Nanomedicine, Journal of Bionanoscience, 8, 137-140 (2014).

[3] M. A. Khanday and A. Najar, Maclaurin's Series Approach for the Analytical Solution of Oxygen Transport to the Biological Tissues Through Capillary Bed, Journal of Medical Imaging and Health Informatics, 5, 959-963 (2015).

[4] H. Huang and F. Yang, An Interpretation of Erlang into Value-passing Calculus, Journal of Networks, 8(7), 1504-1513 (2013).

[5] H. B. Kekre and T. K. Sarode, Vector Quantized Codebook Optimization Using Modified Genetic Algorithm, IETE Journal of Research, 56(5), 257-264 (2010).

[6] M. Gera, R. Kumar, V. K. Jain, Fabrication of a Pocket Friendly, Reusable Water Purifier Using Silver Nano Embedded Porous Concrete Pebbles Based on Green Technology, Journal of Bionanoscience, 8, 10-15 (2014).

[7] M. S. Kumar and S. N. Devi, Sparse Code Shrinkage Based ECG De-Noising in Empirical Mode Decomposition Domain, Journal of Medical Imaging and Health Informatics, 5, 1053-1058 (2015).

[8] R. Bhadada and K. L. Sharma, Evaluation and Analysis of Buffer Requirements for Streamed Video Data in Video on Demand Applications, IETE Journal of Research, 56(5), 242-248 (2010).

[9] M. Kurhekar and U. Deshpande, Deterministic Modeling of Biological Systems with Geometry with an Application to Modeling of Intestinal Crypts, Journal of Medical Imaging and Health Informatics, 5, 1116-1120 (2015).

[10] C. Zhou, Y. Li, Q. Zhang and B. Wang, An Improved Genetic Algorithm for DNA Motif Discovery with Gibbs Sampling Algorithm, Journal of Bionanoscience, 8, 219-225 (2014).

[11] K. Aleksandrovskaya, and VP. Urakhchin, Prediction of the displacements of concrete gravity dams on rock foundations, Power Technology and Engineering (formerly Hydrotechnical Construction) 8, 419-427 (1974).

[12] CR. Allen, and LS. Cluff, Active faults in dam foundations: an update, Proc. 12th World Conf. on Earthquake Engineering, Auckland, New Zealand (2000).

[13] M. Hosseinzadeh, The Effect of Polyurethane Hardness on the New Die-Set of Sheet Hydroforming Process Parameters, Applied Mechanics and Materials, 152: 1623-1627 (2012).

[14] S. A. Zamani, and M. Hosseinzadeh, Investigation of the Effect of Blank Shape on the Sheet Hydroforming Parameters, Applied Mechanics and Materials, 152: 1658-1663 (2012).

[15] M. Wieland, R.P. Brenner, and P. Sommer, Earthquake resilience of large concrete dams: Damage, repair, and strengthening concepts, Trans. 21st Int. Congress on Large Dams, Montreal 131-150 (2003).

[16] J. Mata, and E. Portela, Application of neural networks to dam safety control, 5the International Conference on Dam Engineering, Lisbon, Portugal (2007).

[17] M. Wieland, and P. Brenner, Potentially active faults in the foundations of large dams part I: Vulnerability of dams TO seismic movements in dam foundation, Proc. 14th World Conf. On Earthquake Engineering, Beijing, China, (2008).

[18] M. Hosseinzadeh, S. A. Zamani, and A. Taheri, An Optimization of the New Die Design of Sheet Hydroforming by Taguchi Method, Advanced Materials Research, 337: 294-299 (2011).

[19] M. Hosseinzadeh, M. Bakhshi-Jooybari, M. Shakeri, H. Mostajeran, and H. Gorji, Application of Taguchi Method in Optimization of New Hydroforming Die Design, STEEL RESEARCH INTERNATIONAL, 81(9): 647-650, (2010).

[20] YS. Kim, and B. Kim, Prediction of relative crest settlement of concrete-faced rockfill dams analyzed using an artificial neural network model, Computers and Geotechnics 35, 313-322 (2008).

[21] J. Mata, Interpretation of concrete dam behavior with artificial neural network and multiple linear regression models, Engineering Structures 33, 903-910 (2011).

[22] M. Bakhshi, A. Gorji, M. Hosseinzade H, M. Jamshidi, G. Mohammad-Alinejad, A New Die Design for Sheet Hydroforming of Complex Industrial Parts, Advanced Materials Research, 83: 1084-1091, (2010).

[23] F. Chan, Integration of Expert System with Analysis Hierarchy Process for the Design of Material Handling Selection System, The Journal of Material Processing Technology 32, 137-145 (2001).

[24] O. Kulak, Multi-Attribute Material Handling Equipment Selection Using Information Axiom, the Third International Conference on Axiomatic Design, Seoul 24-24 (2004).

[25] M. hosseinzadeh, H. Mostajeran, M. Bakhshi, A. H. Gorji, S. Norouzi, S. J. Hosseinipour, Experimental and Numerical Investigations of the Effective Parameters in a New Sheet Hydroforming Method, AEROSPACE MECHANICS JOURNAL, 5: 55-68 (2009).

[26] M. Hosseinzadeh, H. Mostajeran, M. Bakhshi-Jooybari, A.H. Gorji, S. Nourouzi, and S. Hosseinipour, Novel combined standard hydromechanical sheet hydroforming process, Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 224(3): 447-457 (2010).

[27] H. Aghajani, M. Elyasi, and M. Hoseinzadeh, Feasibility study on aluminum alloys and A441 AISI steel joints by friction stir welding, International Journal of Advanced Design and Manufacturing Technology, 7(4): 99-109 (2015).

[28] S. Chakraborthy, Design of Material Handling Equipment Selection Model Using Analytic Hierarchy Process, International Journal of Advanced Manufacturing 28, 1237-1245 (2006).

[29] AT. Gumus, Evaluation of Hazardous Waste Transportation Firms by Using a Two Step Fuzzy-AHP and TOPSIS Methodology, Expert Systems with Applications 36, 4067-4074 (2009).

[30] E. Nateghi, and M. Hosseinzadeh, Experimental investigation into effect of cooling of traversed weld nugget on quality of high-density polyethylene joints, The International Journal of Advanced Manufacturing Technology, (2015): 1-14.

[31] H. Naghibi, M. Shakeri, and M. Hosseinzadeh, Neural Network and Genetic Algorithm Based Modeling and Optimization of Tensile Properties in FSW of AA 5052 to AISI 304 Dissimilar Joints, Transactions of the Indian Institute of Metals (2015): 1-10.

[32] M. Elyasi, H. Aghajani, and M. Hoseinzadeh, Study on joint zone Microstructure Evolution and Hardness in Friction Stir Welding of AA1100 Aluminum alloy to A441 AISI steel, Modares Mechanical Engineering, 14(4): 97-107 (2015).

[33] S. A. Zamani, and M. Hosseinzadeh, Using ANFIS by BA optimization for modeling of plastic injection molding process, International Journal of Engineering and Technology, 2(6): 474-486 (2014).

[34] M. Hosseinzade, and A. Zamanib, Experimental and numerical investigation of tube sinking of rectangular tubes from round section, International Journal of Engineering and Technology, 2(5): 407-419 (2014).

[35] A. Bascetin, An Application of the Analytic Hierarchy Process in Equipment Selection at Orhaneli Open Pit Coal Mine, Mining Technology (Trans. Inst. Min. Metall. A) 113, A192-A199 (2004).

[36] G. Enea, and G. Galante, An Integrated Approach to the Facilities and Material Handling System Design, International Journal of Production Research 40, 4007-4017 (2002).

[37] N. Celebi, An Equipment Selection and Cost Analysis System for Open Pit Coal Mines, International Journal of Mining, Reclamation and Environment 12, 181-187 (1998).

[38] S. Iftikhar, F. Ahmad and K. Fatima, A Semantic Methodology for Customized Healthcare Information Provision, Information Sciences Letters, 1(1), 49-59 (2012).

[39] P. D. Sia, Analytical Nano-Modelling for Neuroscience and Cognitive Science, Journal of Bioinformatics and Intelligent Control, 3, 268-272 (2014).

[40] C. Guler, Production of particleboards from licorice (Glycyrrhiza glabra) and European black pine (Pinus Nigra Arnold) wood particles, Scientific Research and Essays, 10(7), 273-278 (2015).

[41] Z. Chen and J. Hu, Learning Algorithm of Neural Networks on Spherical Cap, Journal of Networks, 10(3), 152-158 (2015).

[42] W. Lu, Parameters of Network Traffic Prediction Model Jointly Optimized by Genetic Algorithm, Journal of Networks, 9(3), 695-702 (2014).

[43] K. Boubaker, An attempt to solve neutron transport equation inside supercritical water nuclear reactors using the Boubaker Polynomials Expansion Scheme, International Journal of Physical Sciences, 7(19), 2730-2734 (2012).

[44] K. Abd-Rabou, Fixed Point Results in G-Metric Space, Mathematical Sciences Letters, 3(3), 141-146 (2014).

[45] Binu and M. Selvi, BFC: Bat Algorithm Based Fuzzy Classifier for Medical Data Classification, Journal of Medical Imaging and Health Informatics, 5, 599-606 (2015).

[46] C. Kamath, Analysis of Electroencephalogram Background Activity in Epileptic Patients and Healthy Subjects Using Dispersion Entropy, Journal of Neuroscience and Neuroengineering, 3, 101-110 (2014).

[47] G. Kaur and E. M. Bharti, Securing Multimedia on Hybrid Architecture with Extended Role-Based Access Control, Journal of Bioinformatics and Intelligent Control, 3, 229-233 (2014).

[48] M. Ramalingam and D. Rana, Impact of Nanotechnology in Induced Pluripotent Stem Cells-driven Tissue Engineering and Regenerative Medicine, Journal of Bionanoscience, 9, 13-21 (2015).

[49] D. Denby, and D. Schofield, Application of Expert Systems in Equipment Selection for Surface Design, International Journal of Mining, Reclamation and Environment 4, 165-171 (1990).

[50] M. Hosseinzade, and A. Zamanib, Experimental and numerical investigation of tube sinking of rectangular tubes from round section, International Journal of Engineering and Technology sciences, 2(5): 407-419 (2014).

[51] S. A. Zamani, S. R. Tahmasbpour, B. Asadi, and M. Hosseinzadeh, SA, Numerical Simulation and Plan Stress Analytical Solution of Rotating Disk in High Speed, Indian Journal of Scientific Research, 3(1): 124-136 (2014).

[52] S. R. Tahmasbpour, S. A. Zamani, B. Asadi, and M. Hosseinzadeh, Numerical Simulation and Plane Stress Analytical Solution of two Connected Rotating Disk, Indian J. Sci. Res, 3(1): 115-123 (2014).

[53] C. Krishnamoorthy, K. Rajamani, S. Mekala and S. Rameshkumar, Fertigation through trickle and micro sprinkler on flowering characters in cocoa (Theobroma cacao L.), Scientific Research and Essays, 10(7), 266-272 (2015).

[54] J. Peng, A New Model of Data Protection on Cloud Storage, Journal of Networks, 9(3), 666-671 (2014).

[55] S. Nageswari and V. Suresh Kumar, VCMIPWM Scheme to Reduce the Neutral-Point Voltage Variations in Three-Level NPC Inverter, IETE Journal of Research, 60(6), 396-405 (2014).

[56] X. Ochoa, Connexions: A Social and Successful Anomaly among Learning Object Repositories, Journal of Emerging Technologies in Web Intelligence, 2(1), 11-22 (2010).

[57] Y.S. Chiu, What Can Signed Languages Tell Us About Brain?, Journal of Neuroscience and Neuroengineering, 1, 54-60 (2012).

[58] K. Rezgui, H. Mhiri and K. Ghédira, Theoretical Formulas of Semantic Measure: A Survey, Journal of Emerging Technologies in Web Intelligence, 5(4), 333-342 (2013).

[59] K. Parwez and S. V. Budihal, Carbon Nanotubes Reinforced Hydroxyapatite Composite for Biomedical Application, Journal of Bionanoscience, 8, 61-65 (2014).

[60] H. El-Owaidy, A. Abdeldaim and A. A. El-Deeb, On Some New Retarded Nonlinear Integral Inequalities and Their Applications, Mathematical Sciences Letters, 3(3), 157-164 (2014).

[61] H. Zheng, J. Jie and Y. Zheng, Multi-Swarm Chaotic Particle Swarm Optimization for Protein Folding, Journal of Bionanoscience, 7, 643-648 (2013).

[62] S. V. Sathyanarayana and K. N. Hari Bhat, Novel Scheme for Storage and Transmission of Medical Images with Patient Information Using Elliptic Curve Based Image Encryption Schemes with LSB Based Steganographic Technique, Journal of Medical Imaging and Health Informatics, 2, 15-24 (2012).

[63] Satyendra Sharma and Brahmjit Singh, Field Measurements for Cellular Network Planning, Journal of Bioinformatics and Intelligent Control, 2, 112-118 (2013).

[64] Y. Wang, An Approximate Algorithm for TSP with Four Vertices and Three Lines Inequality, Information Sciences Letters, 3(2), 41-44 (2014).

[65] S. P. Singh and B. K. Konwar, Insilico Proteomics and Genomics Studies on ThyX of Mycobacterium Tuberculosis, Journal of Bioinformatics and Intelligent Control, 2, 11-18 (2013).

[66] K. Kumar, A. K. Verma and R. B. Patel, Framework for Key Management Scheme in Heterogeneous Wireless Sensor Networks, Journal of Emerging Technologies in Web Intelligence, 3(4), 286-296 (2011).

[67] N. Barnthip and A. Muakngam, Preparation of Cellulose Acetate Nanofibers Containing Centella Asiatica Extract by Electrospinning Process as the Prototype of Wound-Healing Materials, Journal of Bionanoscience, 8, 313-318 (2014).

[68] R. Jac Fredo, G. Kavitha and S. Ramakrishnan, Segmentation and Analysis of Corpus Callosum in Autistic MR Brain Images Using Reaction Diffusion Level Sets, Journal of Medical Imaging and Health Informatics, 5, 737-741 (2015).

[69] Wang, B. Zhu, An Improved Algorithm of the Node Localization in Ad Hoc Network, Journal of Networks, 9(3), 549-557 (2014).

[70] T. Buvaneswari and A. A. Iruthayaraj, Secure Discovery Scheme and Minimum Span Verification of Neighbor Locations in Mobile Ad-hoc Networks, Australian Journal of Basic and Applied Sciences, 8(2), 30-36 (2014).

[71] Y. Zhang, Z. Wang and Z. Hu, Nonlinear Electroencephalogram Analysis of Neural Mass Model, Journal of Medical Imaging and Health Informatics, 5, 783-788 (2015).

[72] S. Panwar and N. Nain, A Novel Segmentation Methodology for Cursive Handwritten Documents, IETE Journal of Research, 60(6), 432-439 (2014).

[73] H. Mao, On Applications of Matroids in Class-oriented Concept Lattices, Mathematical Sciences Letters, 3(1), 35-41 (2014).

[74] D. Kumar, K. Singh, V. Verma and H. S. Bhatti, Synthesis and Characterization of Carbon Quantum Dots from Orange Juice, Journal of Bionanoscience, 8, 274-279 (2014).

[75] V. Kumutha and S. Palaniammal, Enhanced Validity for Fuzzy Clustering Using Microarray data, Australian Journal of Basic and Applied Sciences, 8(3), 7-15 (2014).

[76] S. R. Tahmasbpour, S. A. Zamani, B. Asadi, and M. Hosseinzadeh, Numerical Simulation and Plane Stress Analytical Solution of two Connected Rotating Disk, Indian Journal of Scientific Research, 3(1): 115-123 (2014).

[77] S. Bandopadhyay, and P. Venkatasubramanian, Expert Systems as Decision Aid in Surface Mine Equipment Election, International Journal of Mining, Reclamation and Environment 1, 159-165 (1987).

[78] S. Anon, Belt Conveyors for Bulk Materials. Conveyor Equipment Manufacturers Association, 5the International Conference on Dam Engineering, Lisbon (2007).

[79] Y. Chang, Applications of the Extent Analysis Method on Fuzzy AHP, European Journal of Operational Research 95, 649-655 (1996).

[80] M. Hosseinzadeh, The Effect of Polyurethane Hardness on the New Die-Set of Sheet Hydroforming Process Parameters, Applied Mechanics and Materials, 152: 1623-1627 (2012).

[81] S. A. Zamani, and M. Hosseinzadeh, Investigation of the Effect of Blank Shape on the Sheet Hydroforming Parameters, In Applied Mechanics and Materials, 152: 1658-1663 (2012).

[82] O. O. E. Ajibola, O. Ibidapo-Obe, and V. O. S. OIunloyo, A Model for the Management of Gait Syndrome in Huntington's Disease Patient, Journal of Bioinformatics and Intelligent Control, 3, 15-22 (2014).

[83] L. Z. Pei, T. Wei, N. Lin and Z. Y. Cai, Electrochemical Sensing of Histidine Based on the Copper Germanate Nanowires Modified Electrode, Journal of Bionanoscience, 9, 161-165 (2015).

[84] M. K. Elboree, Explicit Analytic Solution for the Nonlinear Evolution Equations using the Simplest Equation Method, Mathematical Sciences Letters, 3(1), 59-63 (2014).

[85] R. Yousef and T. Almarabeh, An enhanced requirements elicitation framework based on business process models, Scientific Research and Essays, 10(7), 279-286 (2015).

[86] K. Manimekalai and M.S. Vijaya, Taxonomic Classification of Plant Species Using Support Vector Machine, Journal of Bioinformatics and Intelligent Control, 3, 65-71 (2014).

[87] S. Rajalaxmi and S. Nirmala, Automated Endo Fitting Curve for Initialization of Segmentation Based on Chan Vese Model, Journal of Medical Imaging and Health Informatics, 5, 572-580 (2015).

[88] T. Mahmood and K. Hayat, Characterizations of Hemi-Rings by their Bipolar-Valued Fuzzy h-Ideals, Information Sciences Letters, 4(2), 51-59 (2015).

[89] Agarwal and N. Mittal, Semantic Feature Clustering for Sentiment Analysis of English Reviews, IETE Journal of Research, 60(6), 414-422 (2014).

[90] S. Radharani and M. L. Valarmathi, Content Based Watermarking Techniques using HSV and Fractal Dimension in Transform Domain, Australian Journal of Basic and Applied Sciences, 8(3), 112-119 (2014).

[91] S. A. Zamani, and M. Hosseinzadeh, Investigation of the Effect of Blank Shape on the Sheet Hydroforming Parameters, Applied Mechanics and Materials, 152: 1658-1663 (2012).

[92] L. Zare, and M. Ataei, The Application of Fuzzy Analytic Hierarchy Process (FAHP) Approach to Selection of Optimum Underground Mining Method for Jajarm Bauxite Mine, Expert Systems With Applications 36, 8218-8226 (2009).

[93] R. Zhang, Y. Bai, C. Wang and W. Ma, Surfactant-Dispersed Multi-Walled Carbon Nanotubes: Interaction and Antibacterial Activity, Journal of Bionanoscience, 8, 176-182 (2014).

[94] B. K. Singh, Generalized Semi-bent and Partially Bent Boolean Functions, Mathematical Sciences Letters, 3(1), 21-29 (2014).

[95] S. K. Singla and V. Singh, Design of a Microcontroller Based Temperature and Humidity Controller for Infant Incubator, Journal of Medical Imaging and Health Informatics, 5, 704-708 (2015).

[96] N. Barnthip and A. Muakngam, Preparation of Cellulose Acetate Nanofibers Containing Centella Asiatica Extract by Electrospinning Process as the Prototype of Wound-Healing Materials, Journal of Bionanoscience, 8, 313-318 (2014).

[97] R. Jac Fredo, G. Kavitha and S. Ramakrishnan, Segmentation and Analysis of Corpus Callosum in Autistic MR Brain Images Using Reaction Diffusion Level Sets, Journal of Medical Imaging and Health Informatics, 5, 737-741 (2015).

[98] Wang, B. Zhu, An Improved Algorithm of the Node Localization in Ad Hoc Network, Journal of Networks, 9(3), 549-557 (2014).

[99] T. Buvaneswari and A. A. Iruthayaraj, Secure Discovery Scheme and Minimum Span Verification of Neighbor Locations in Mobile Ad-hoc Networks, Australian Journal of Basic and Applied Sciences, 8(2), 30-36 (2014).

[100] Y. Zhang, Z. Wang and Z. Hu, Nonlinear Electroencephalogram Analysis of Neural Mass Model, Journal of Medical Imaging and Health Informatics, 5, 783-788 (2015).

[101] S. Panwar and N. Nain, A Novel Segmentation Methodology for Cursive Handwritten Documents, IETE Journal of Research, 60(6), 432-439 (2014).

[102] H. Mao, On Applications of Matroids in Class-oriented Concept Lattices, Mathematical Sciences Letters, 3(1), 35-41 (2014).

[103] D. Kumar, K. Singh, V. Verma and H. S. Bhatti, Synthesis and Characterization of Carbon Quantum Dots from Orange Juice, Journal of Bionanoscience, 8, 274-279 (2014).

[104] V. Kumutha and S. Palaniammal, Enhanced Validity for Fuzzy Clustering Using Microarray data, Australian Journal of Basic and Applied Sciences, 8(3), 7-15 (2014).

[105] Y. Wang, C. Yang and J. Yu, Visualization Study on Cardiac Mapping: Analysis of Isopotential Map and Isochron Map, Journal of Medical Imaging and Health Informatics, 5, 814-818 (2015).

[106] R. Su, Identification Method of Sports Throwing Force Based on Fuzzy Neural Network, Journal of Networks, 8(7), 1574-1581 (2013).

[107] L. Feng, W. Zhiqi and L. Ming, the Auto-Detection and Diagnose of the Mobile Electrocardiogram, Journal of Medical Imaging and Health Informatics, 5, 841-847 (2015).

[108] D. Deepa, C. Poongodi, A. Bharathi, Satellite Image Enhancement using Discrete Wavelet Transform and Morphological Filtering for Remote Monitoring, Australian Journal of Basic and Applied Sciences, 8(3), 27-34, (2014).

[109] R. Sarma, J. M. Rao and S. S. Rao, Fixed Point Theorems in Dislocated Quasi-Metric Spaces, Mathematical Sciences Letters, 3(1), 49-52 (2014).

[110] T. Gopalakrishnan and P. Sengottvelan, Discovering user profiles for web personalization using EM with Bayesian Classification, Australian Journal of Basic and Applied Sciences, 8(3), 53-60 (2014).

[111] R. Uthayakumar and A. Gowrisankar, Generalized Fractal Dimensions in Image Thresholding Technique, *Inf. Sci. Lett.* 3(3), 125-134 (2014).

[112] B. Ould Bilal, D. Nourou, C. M. F Kébé, V. Sambou, P. A. Ndiaye and M. Ndongo, Multi-objective optimization of hybrid PV/wind/diesel/battery systems for decentralized application by minimizing the levelized cost of energy and the CO2 emissions, *International Journal of Physical Sciences,* 10(5), 192-203 (2015).

[113] A. Maqbool, H. U. Dar, M. Ahmad, G. N. Malik, G. Zaffar, S. A. Mir and M. A. Mir, Comparative performance of some bivoltine silkworm (Bombyx mori L.) genotypes during different seasons, *Scientific Research and Essays,* 10(12), 407-410 (2015).

[114] M N. Seyedaghaee, S. Amirgholipour, H. Alinejad-Rokny and F. Rouhi, Strategic Planning Formulation by Using Reinforcement Learning, *Research Journal of Applied Sciences, Engineering and Technology*, 4(11), 1448-1454 (2012).

[115] S. Y. Li, C. Chen, H. Q. Zhang, H. Y. Guo, H. Wang, L. Wang, X. Zhang, S. N. Hua, J. Yu, P. G. Xiao, R. S. Li, and X. Tan, *Antiviral. Res*. 67, 18 (2005).

[116] R. B. Little, A theory of the relativistic fermionic spinrevorbital, *International Journal of Physical Sciences,* 10(1), 1-37 (2015).

[117] Z. Chen, F. Wang and Li Zhu, The Effects of Hypoxia on Uptake of Positively Charged Nanoparticles by Tumor Cells, *J. Bionanosci.* 7, 601-605 (2013).

[118] A. Kaur and V. Gupta, A Survey on Sentiment Analysis and Opinion Mining Techniques, *Journal of Emerging Technologies in Web Intelligence*, 5(4), 367-371 (2013).

[119] P. Saxena and M. Agarwal, Finite Element Analysis of Convective Flow through Porous Medium with Variable Suction, *Inf. Sci. Lett.* 3(3), 97-101 (2014).

[120] J. G. Bruno, Electrophoretic Characterization of DNA Oligonucleotide-PAMAM Dendrimer Covalent and Noncovalent Conjugates, *J. Bionanosci.* 9, 203-208 (2015).

[121] K. K. Tanaeva, Yu. V. Dobryakova, and V. A. Dubynin, Maternal Behavior: A Novel Experimental Approach and Detailed Statistical Analysis, *J. Neurosci. Neuroeng.* 3, 52-61 (2014).

[122] Zaitseva and M. Rusin, Healthcare System Representation and Estimation Based on Viewpoint of Reliability Analysis, *J. Med. Imaging Health Inf.* 2, 80-86 (2012).

[123] R. Ahirwar, P. Devi and R. Gupta, Seasonal incidence of major insect-pests and their biocontrol agents of soybean crop (Glycine max L. Merrill), *Scientific Research and Essays*, 10(12), 402-406 (2015).

[124] Boussak, H. Chemani and A. Serier, Characterization of porcelain tableware formulation containing bentonite clay, *International Journal of Physical Sciences,* 10(1), 38-45 (2015).

[125] Q. Xiaohong, and Q. Xiaohui, An Evolutionary Particle Swarm Optimizer Based on Fractal Brownian Motion. *Journal of Computational Intelligence and Electronic Systems*, 1, 138 (2012).

[126] Minhas and M. Kumar, LSI Based Relevance Computation for Topical Web Crawler, *Journal of Emerging Technologies in Web Intelligence*, 5(4), 401-406 (2013).

[127] H. Alinejad-Rokny, H. Pourshaban, A. Goran, M. MirnabiBaboli, Network Motifs Detection Strategies and Using for Bioinformatic Networks, *Journal of Bionanoscience,* 8, 353 (2014).

[128] Y. Shang, Efficient strategies for attack via partial information in scale-free networks, *Inf. Sci. Lett.* 1(1), 1-5 (2012).

[129] M. Ahmadinia, H. Alinejad-Rokny, and H. Ahangarikiasari, *Journal of Networks* 9(10), 2567-2573 (2014).

[130] I. Rathore and J. C. Tarafdar, Perspectives of Biosynthesized Magnesium Nanoparticles in Foliar Application of Wheat Plant. *J. Bionanosci.* 9, 209-214 (2015).

[131] Hui Yan and Haiyan Hu, Research and Realization of ISIC-CDIO Teaching Experimental System Based on RFID Technology of Web of Things, *J. Bionanosci.* 7, 696-702 (2013).

[132] R. Teles, B. Barroso, A. Guimaraes and H. Macedo, Automatic Generation of Human-like Route Descriptions: A Corpus-driven Approach, *Journal of Emerging Technologies in Web Intelligence*, 5(4), 413-423 (2013).

[133] E. S. Hui, Diffusion Magnetic Resonance Imaging of Ischemic Stroke, *J. Neurosci. Neuroeng.* 1, 48-53 (2012).

[134] O. E. Emam, M. El-Araby and M. A. Belal, On Rough Multi-Level Linear Programming Problem, *Inf. Sci. Lett.* 4(1), 41-49 (2015).

[135] B. Prasad, D.C. Dimri and L. Bora, Effect of pre-harvest foliar spray of calcium and potassium on fruit quality of Pear cv. Pathernakh, *Scientific Research and Essays,* 10(11), 392-396 (2015).

[136] L. Xu, A. Y. Hui, E. Albanis, M. J. Arthur, S. M. O'Byrne, W. S. Blaner, P. Mukherjee, S. L. Friedman, and F. J. Eng, *Gut*. 54, 142 (2005).

[137] G. Singh, Optimization of Spectrum Management Issues for Cognitive Radio, *Journal of Emerging Technologies in Web Intelligence*, 3(4), 263-267 (2011).

[138] D. Madhuri, Linear Fractional Time Minimizing Transportation Problem with Impurities, *Inf. Sci. Lett.* 1(1), 7-19 (2012).

[139] A. D. Dele, Oscilating magnetic field an anti malaria therapy, *International Journal of Physical Sciences,* 10(10), 329-334 (2015).

[140] Mayevsky, J. Sonn and E. Barbiro-Michaely, Physiological Mapping of Brain Functions In Vivo: Surface Monitoring of Hemodynamic Metabolic Ionic and Electrical Activities in Real-Time, *J. Neurosci. Neuroeng.* 2, 150-177 (2013).