International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 11 • Issue 1 pp. 503-510 Jan 2019-June 2019 www.csjournals.com

Two-Stage Stochastic Algorithm for View Maintenance model: Cost-Efficiency and Maintenance Optimization

Rolly Gupta[a], Sangeeta Sabharwal [b], Anjana Gosain [c]
[a]Research Scholar, Department of COE, Netaji Subhas Institute of Technology, Delhi University, Delhi, India
[b]Professor, Department of COE, Netaji Subhas Institute of Technology, Delhi University, Delhi, India
[c]Professor, University School of Information Technology,GGS Indraprastha University, Delhi, India
[a]rollygupta02@gmail.com, [b]ssab63@gmail.com, [c]anjana_gosain@hotmail.com

Abstract: This paper develops a two-stage stochastic cost-efficiency optimized View Maintenance (VM) model subjected to uncertain schema changes, which leads to uncertain views and model refreshment. Optimization of VM models are used for making decisions of model configurations to be applied in different and occasional diverse future uncertain scenarios. Views efficiency and cost data can provide the necessary information of conditions and inputs to determine an optimum preventive maintenance policy. In this formulation, the VM model is exposed to distinct usage scenarios that are collectively represented as the future usage profile. This profile is described by a set of specific scenarios and their probability of occurrence or likelihood. Once these future usage profiles are known or can be estimated, the cost rate for prospective VM model configurations and associated maintenance can be modelled. The cost-efficiency model design with maintenance modelling problem is defined as a two-stage stochastic cost-efficiency programming problem with recourse. The decision variables for the first-stage are the selection and the number of views to be used in the VM model, where the cost rate objective function is constrained by views availability. The second stage variables are defined by the corrective and preventive maintenance plan in order to optimize the maintenance time interval for planned refreshment of views in the VM model. A numerical example demonstrates that the VM model cost rate can be optimized by these decision variables.

Keywords: Two-stage stochastic problem, Preventive maintenance, Efficiency

Introduction

There have been significant researches related to optimization of View maintenance (VM) model efficiency based on long term forecasts of usage requirements and schema changes. In any real application, the efficiency of VM model and its views naturally changes due to schema changes and data changes. The schema changes and data changes affecting views efficiency then influence VM model efficiency design. The new efficiency analysis and optimization of VM model involving cost minimization are evaluated based on variations of schema under changing circumstances. Preventive maintenance (PM) optimization is conducted simultaneously with the VM model efficiency. In the VM modelling, the model is exposed to distinct usages, which are represented as the future usage profile.

In efficiency optimization, redundancy allocation is a technique to increase VM model efficiency for configuration design. Redundancy allocation problems with multiple choices of views are NP-hard. Several techniques have been studied to solve efficiency optimization for redundancy allocation problems. Heuristics, meta-heuristics or exact algorithms have been used for solving redundancy allocation problems [1]. In addition, there are numerous techniques for optimization of PM policy. Maintenance optimization is formulated to minimize costs of maintenance and/or to increase model availability [2,3]. Although many researches focus on PM policy, the best method to solve the PM optimization problem is still not always understood.

The PM optimization for model was studied to minimize cost per unit time [4]. The optimization model of [5] is solved by Bender's decomposition to schedule PM considering the total cost objective function. The PM scheduling models have also been formulated as mixed-integer linear programming VM models [5,6]. Cheng et al [7] presented optimization models to minimize model cost subject to efficiency considering the PM time interval. The analytical model of PM was analyzed by considering the expected cost rate objective function and PM policy based on a Weibull distribution in [8]. They develop a mathematical formulation of PM period and the number of PM events.

International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 11 • Issue 1 pp. 503-510 Jan 2019-June 2019 www.csjournals.com

When uncertainty lies in views or VM model efficiency for different scenarios, changes of schema, usage requirements, or data change can cause different efficiency characteristics for the VM model. Efficiency optimization research has been developed for considering uncertainty using mathematical programming the VM models. In general, stochastic programming (SP) with recourse to consider uncertainty has been studied. Dantzig [9], Birge and Louveaux [10] and Kall and Wallace [11] proposed the formulation of stochastic problems with recourse and reviewed the basic concepts, methods and applications. The mathematical program where parameters in the objective function or constraints are uncertain is considered as stochastic programming. When some parameters are random, then solutions and the optimal objective value of the problem are also random. The two-stage stochastic decision is for dealing with uncertainty. After the first-stage decision, a random event occurs affecting the outcome. Solving a recourse problem is to evaluate the second stage that compensates for any impact that might have been experienced from the first-stage decision. The two-stage stochastic programming model has been formulated in many standard forms. Two-stage stochastic integer programming and two-stage stochastic mixed integer programmings have been developed [12].

In this paper, maintenance costs and VM model efficiency improvement for designs are studied and the two-stage stochastic programming problem that relates VM model efficiency optimization and PM is formulated. Many objective frameworks with randomness of future usages are proposed for the two-stage stochastic problem. The decision variable in first-stage is the number of views used in each VM model affecting VM model cost-rate and VM model efficiency. The decision makers must make a decision at the current time from the best-known information. For the second-stage decision, corrective actions are designed based on realization of full information. The corrective actions represent the maintenance time interval for refreshment of the views in the VM model.

Future usages Profile of Data changes and Schema changes

Uncertainty in future usages can involve several factors of uncertain schema/load to views within the future usage profile, as shown in Figure 1. Sometimes, VM model efficiency estimation is often problematic due to unplanned variation, or changing schema changes. More often, not only data changes and schema changes but also business requirements changes can be uncertain and dependent on usage conditions of views and models in the future.

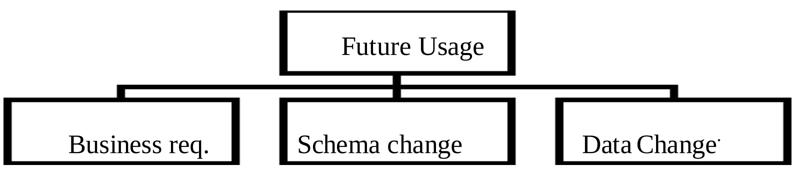
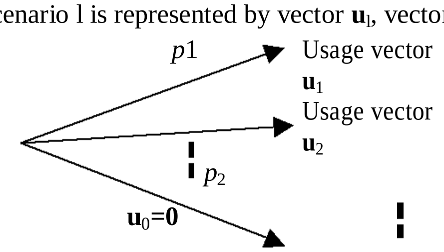


Figure 1. Uncertainty in Future usages

Views/schema parameters are represented in the future usage scenarios as vectors, which describe how changing schema changes in different future usage scenarios affect efficiency functions. Uncertain schema factors are demonstrated in the form of a random vector **U** in the future usage profile. The VM model parameter U_k is a random variable of the k^{th} schema factor that the views or VM model experiences. A random future usage vector is represented with c different operating usage and schema factors where each uncertain usage condition has different effects on each views, $\mathbf{U} = (U_1, U_2, \ldots, U_c)$.

The future usage profile defines possible future usage scenarios defined by future usage schema vectors. In practice, the possible future usage scenarios can be enumerated from prediction of how the VM model will be used and possible occurrences of each future usage. The future usage schema vector in each scenario l is determined from random future usage vector **U**. There are c different factors of future usage schema changes which determine how schema changes of different factors change in that future usage scenario. For example, the determined future usage vector for possible future scenario l is represented by vector $\mathbf{u}_l$, vector $\mathbf{u}_2$ for possible future scenario 2 and so on.



p_v Usage vector $\mathbf{u}_v$

Figure 2. Discrete scenarios in future usage profile

In Figure 2, the possible future usage vectors in the future usage profile are defined as $\mathbf{u}_l$ and each future usage scenario is associated with a probability p_l. Let $\mathbf{u}_l = (u_{1l}, u_{2l},\ldots,u_{cl})$ represent a determined vector of data change and/or schema changes. For example, for electronic, temperature is a critical contributor to views failure (u_{1l}), and for a given future usage scenario, the risk of views failure increases along with increasing temperatures. For mechanical, mechanical loading (u_{2l}) and schema (u_{3l}) are important factors. The current data and schema vector is $\mathbf{u}_0$ and it is known with certainty and given as $\mathbf{u}_0 = \mathbf{0}$. Each schema from future usage profile is defined by the usage vector $\mathbf{u}_l$. All schema changes have been scaled from 0 to 1. For example, u_{2l} is a second usage variable at future l.

View Maintenance Model

For this paper, a parallel VM model maintenance is introduced for the stochastic efficiency model, as shown in Figure 3. Since the number of multiple views, x, is connected within a VM model, the VM model efficiency is given in Eq. (1).

$$R(x;t) = 1-(1-r)^x \tag{1}$$

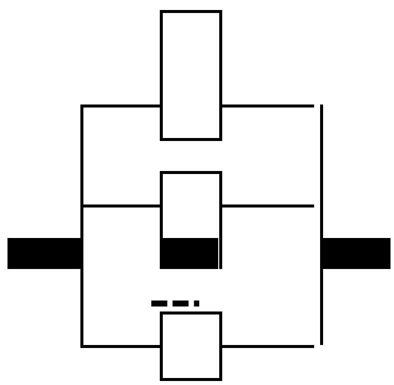


Figure 3. Parallel VM model

For our proposed VM model efficiency function, the future usage profile of data changes and schema changes is considered for developing a model efficiency. The views efficiency (r_i) is no longer constant, but a random variable. The number of multiple views, x, is connected and the VM model efficiency is given as follows:

$$R(x, \mathbf{U};t) = 1-(1-r(\mathbf{U};t))^x \tag{2}$$

$\mathbf{U}$ is a random future usage vector, so $r(\mathbf{U};t)$ and $R(x,\mathbf{U};t)$ are also random variables. The VM model efficiency equation is derived from the views reliabilities based on the future usage profile.

A. Maintenance Policy with refreshment time interval variable

A time refreshment maintenance policy of stochastic VM model efficiency is modelled from a renewal reward process that contains multiple random variables.. The refreshment for the entire VM model is applied after the time interval y. When the VM model is outdated before time y, the refreshment is done immediately. Long run average time refreshment cost per unit time can be evaluated as:

$$\lim_{t\to\infty}(C(t)/t) = \frac{\text{Expected maintenance cost between the two refreshments}}{\text{Expected time between the two refreshments}} = \frac{E(TC)}{E(U)} \tag{3}$$

where TC is the total cost of a renewal cycle, and U is the length of the renewal cycle. The cycle time is equal to time y for preventive refreshment. Then the expected total maintenance cost for a VM model is given as: (4)

$$E(TC) = (C_F + C_R)(1 - R(y)) + C_R R(y) = C_F (1 - (y)) + C_R$$

In each cycle, we have one refreshment, and the cost C_R should be considered in all cases. However, if the VM model is outdated before the fixed refreshment interval, C_F, the cost on an unanticipated outdation, should also be included, and the probability of outdation before refreshment interval y is 1 - $R(y)$. Expected time between two refreshments is

$$E[U] = yR(y) + \int_0^y tf(t)\,dt = \int_0^y R(t)\,dt\,. \tag{5}$$

Based on Equations (4) to (5), the avertime long-run maintenance cost rate is given as (6)

$$CR(y) = \frac{C_F (1 - R(y)) + C_R}{yR(y) + \int_0^y tf(t)\,dt} = \frac{C_F (1 - R(y)) + C_R}{\int_0^y R(t)\,dt}\,.$$

Notation

U	=	random schema vector, $\mathbf{U} = (U_1, U_2, \ldots, U_c)$
$\mathbf{u}l$	=	Random schema variable **U** $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_v\}$
C	=	cost constraint
α_{ik}	=	Sensitivity coefficient for views i and schema factor k
x	=	Design variable of number of viewss
y	=	Maintenance inspection interval variable in different scenarios (y_l)
$C(t)$	=	cumulative maintenance cost by time t
$CR(y)$	=	average time long-run maintenance cost rate
$E[U]$	=	expected value of the renewal cycle length
$E[TC]$	=	expected value of the total maintenance cost of the renewal cycle
C_R	=	refreshment cost per unit
C_F	=	cost of refreshment caused by outdation

Mathematical Formulation

To illustrate the proposed VM modelling, the two-stage stochastic problem is presented to model design efficiency under uncertainty. The long-term forecast is used for making decisions about the VM model configuration to be implemented in different future usage scenarios. However, the future conditions cannot be explicitly predicted. The decision makers must decide without fully understanding of future information. The decision variables in the first-time are represented as the number of views composed in the VM model. The second-time decision variables represent when the preventive or corrective actions are taken. Decision variables of maintenance refreshment time interval are determined to optimize the VM model maintenance cost rate.

The number of views and maintenance time interval are decision variables in the objective function. Future usages depending on schema factors are also the parameters in the VM model and considered as random variable. This two -stage stochastic programming with recourse represents the cost-rate model design with maintenance modelling in Eq. (7).

$$\min c_A(x) + Q(x, \mathbf{y})$$
$$\text{s.t.} \quad 1 \le x \le x_{max}$$
$$x \quad \{0,1,2,3,...\}$$

where

$$Q(x, \mathbf{y}) = E_{\mathbf{U}}[Q(\mathbf{U}, \mathbf{y}; x)] = \sum_{l=1}^{v} p_l Q(\mathbf{u}_l, \mathbf{y}; x)$$

$$c_A(x) = cx(A/P, int\%, N)$$

Reliability of system: $R(x,\mathbf{y},\mathbf{U}) = 1 - (1 - r(y,\mathbf{U}))x$

Component reliability: $r(\mathbf{y},\mathbf{U}) = \exp(-(\mathbf{y}/\eta(\mathbf{U}))\beta)$ (7) & (8)

The net present value formula $(A/P, int, N) = int(1+int)N / (1+int)N - 1$ is the relationship between P and A for finding a uniform series of end-of-period cash flows of amount A for N periods. c_A(x) is the equivalent end of period cash flows in a uniform series for a specified number of periods, starting at the end of the first period and continuing through the last period.

The optimization VM model given in (7) is known as implicit representation of the stochastic program. By considering the value function or recourse function in (8), the condense representation can be shown in Eq. (9)

$$\min cx(A\,P, int\%, N) + \sum_{l=1}^{v} pl\, Q(\mathbf{u}l, \mathbf{y};\, x)$$

$$\text{s.t.} \quad 1 \le x \le x_{max} \tag{9}$$
$$x \ \varepsilon \ \{0,1,2,3,...\}$$

where

$$Q(yl, \mathbf{U}; \mathbf{x}) = CF\,[(1 - r(yl, \mathbf{U}))x] + CR/ o\text{-}y \int R(t; \mathbf{U}, x)dt$$

In the VM model, the decision variable, x, is required to have integer solution while another decision variable, y_l, is continuous. The technique presented to solve the two-time stochastic VM model is a heuristic approach, which is used to obtain integer solutions. In order to simplify computational complexity from mixed-integer programming, one decision variable is searched along with another to optimize the best solutions.

A. Numerical Example

The optimization VM model of VM model cost-efficiency and maintenance are designed based on parameter values given in [15-18]. Future usage profile is expressed in Fig. 4 with three future usage schema variables and there are four probabilities p_l = (0.4 0.4 0.1 0.1) for four scenarios. The future usage schema variables are related to changes of schema, data changes, and requirement changes have been scaled from zero to one. Each possible scenario is used for the decisions corresponding to each possible future usage.

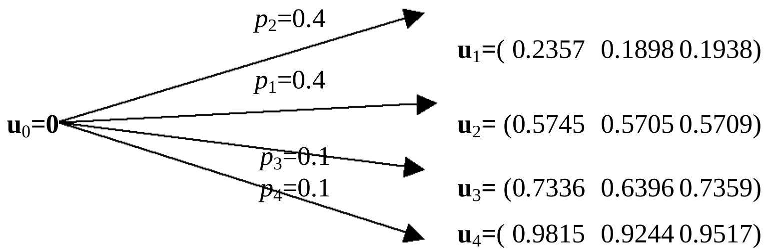


Figure 4. Future usage probability ($\mathbf{u}_l$ is the operating usage and schema vector, l=1,2,…v)

The future usage profile displays the probabilities associated with data and schema $\mathbf{u}_l$ vector. Each future usage vector represents one scenario that the views in the VM model may experience. The current data and usage schema vector is given as $\mathbf{u}_0 = \mathbf{0}$.

An example problem is solved to demonstrate the numerical results by using the optimization technique for the optimal continuous solution. Neighborhood search algorithm is performed to solve the mathematical formulation for the recommended integer solution.

Consider a parallel VM model where number of views represented by decision variable x. Views cost is c. The net present value and equivalent cash flow annual cost method are used to evaluate the annual VM model cost and budget with interest rate (int = 0.2) and period of time (N =20).

The decision makers make a decision at the current time from the best known information in the future usage profile. Given the four future usage scenarios, the second-stage or correction action can be computed from $\sum p_l\, Q(\mathbf{u}_l, \mathbf{y}; x)$. The maintenance time interval is represented by decision variable y_l. The pattern search algorithm and combinatorial neighbourhood search are used to find the recommended solutions of the number of views and the optimal maintenance policy.

This example demonstrates the VM model feasibility and the solutions of two-stage stochastic VM model are shown in Table 1. The recommended number that optimizes annual VM model cost and maintenance is not more than 3 multiple views in parallel. The maintenance refreshment intervals from the second-stage decision are taken for each scenario and also shown in Table 1.

Table 1**.** Solution of parallel VM model maintenance

		Min Value	Max value
	Number	1	3
Series 1	Maintenance time interval	1	–
Series 2	Maintenance time interval	2	–
Series 3	Maintenance time interval	22	–
Overall	Annual VM model cost ($/unit time)	9501	–

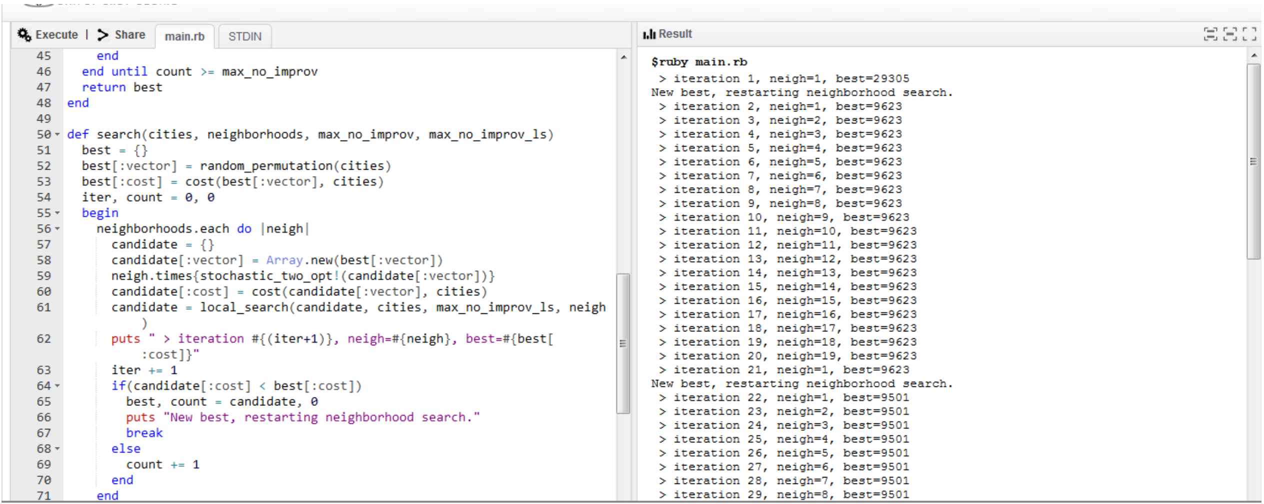


Figure 5: Implementation code

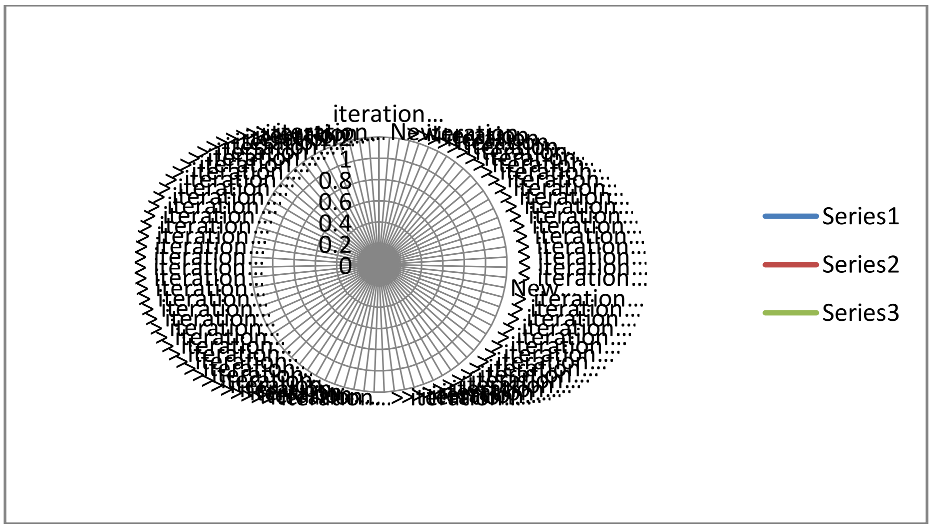


Figure 6: Result

Conclusion

Maintenance costs and VM model efficiency improvement for designs are studied and the two-stage stochastic programming problem that relates VM model efficiency optimization and PM is formulated. The decision variable in first-stage is the number of views used in each VM model affecting VM model cost-rate and VM model efficiency For the second-stage decision, corrective actions are designed based on realization of full information. The corrective actions represent the maintenance time interval for refreshment of the views in the VM model.

The study of two-stage stochastic problem is presented due to uncertainty occurring in the different future usage scenarios. The mathematical VM model of two-stage stochastic program considering future usage profile is demonstrated and provides recommend solutions for VM model configuration and updated maintenance policy in the numerical example.

References

1. Kuo W., Zuo Ming J., Optimal Efficiency VM modeling: Principles and Applications, Wiley, 2003
2. Vassiliadis C. G., Pistikopoulos E. N., "Chemical-process design and maintenance optimization under uncertainty: A simultaneous approach," *Proc. Ann. Efficiency and maintainability Symp*., 1999 Jan, pp 78 - 83.
3. Vatn J., P. Hqkstad and L. Bodsberg, “An overall model for maintenance optimization”, *Efficiency Engineering and VM model Safety*, vol. 51, 1996, pp. 241-257.

4. Zijlstra, M., Optimization of preventive maintenance for two-unit cold standby redundant models with a general cost structure, *Zeitschrift für Operations Research* vol. 27 issue 1 December 1983. p. 55 - 63
5. Budai, G., Huisman, D., Dekker, R., (2006) Scheduling preventive railway maintenance activities, *Journal of the Operational Research Society*, v 57, n 9, Sep 2006, p 1035-44.
6. Canto, S.V., (2006) Application of Benders' decomposition to power plant preventive maintenance scheduling, *European Journal of Operational Research*, v184, Issue 2, Jan 2008, p 759-777.
7. Cheng, C.Y., Chen, M., Guo, R., (2007) The optimal periodic preventive maintenance policy with degradation rate reduction under efficiency limit, 2007 *IEEE International Conference* on IE and EngMantimement, 2-4 December 2007, Singapore, Singapore, p 649-653.
8. Lim, J.H., Park, D. H., (2007) Optimal periodic preventive maintenance schedules with improvement factors depending on number of preventive maintenances, *Asia-Pacific Journal of Operational Research,* v 24, n 1, February 2007, p 111-124.
9. Dantzig GB , Linear programming under uncertainty. Mngt Sci **1**: 197–206, 1995
10. Birge JR and Louveaux F, Introduction to Stochastic Programming.Sp ringer: New York, 1997
11. Kall P and Wallace SW, Stochastic Programming.John Wiley & Sons: Chichester, 1995
12. Ahmed S., Two-stage stochastic integer programming: A brief introduction.
13. Ahmed S., Review of Introduction to Stochastic Programming by J. R. Birge and F. *Louveaux,Interfaces*, vol. 28, pp. 79, 1998.
14. Chatwattanasiri N., Coit D., Rodgers M. & Song S., Model Efficiency Optimization Considering Uncertain Future Data changes and Usage Schema changes, *Proc of the 18th ISSAT International Conf on Reli and Quality in Design*, Boston, MA, Jul 2012
15. Coit, D., Jin, T., Wattanapongsakorn, N., "Model Optimization Considering Views Efficiency Estimation Uncertainty: A Multi-Criteria Approach," *IEEE Trans on Rel*, 53(3), 369-380, 2004.
16. D. E. Goldberg, "Genetic algorithms in search, optimizationand machine learning", *Addison Wesley,* Reading(MA), 1989.
17. M. Gregory, "Genetic algorithm optimisation of distributed database queries", *Proc. of ICEC'98,* pages 271-276,1998.
18. H. Gupta, "Selection of views to materialize in a data warehouse", Proceedings of the International Conference on Data Engineering, Burmingham, *U.K.,* pages 98-112, April, 1997.
19. H. Gupta and I. S. Mumick, "Selection of views to materialize under a maintenance cost constraint", Proceedings of the International Conference on Data Engineering, 1998.
20. W. J. Labio, D. Quass, and B. Adelberg, "Physical database design for data warehouses", Proceedings of the International Conference on Data Engineering, pages 277-288, 1997.
21. S. Timos, S. Kyuseok and D. Nau, "Improvements on a heuristic algorithm for multiplequery optimization", Data and Knowledge Engineering, 12:197-222, 1994.
22. D. Theodoratos and T. Sellis, "Data warehouse configuration", Proceedings of the 23rd VLDB Conference Athens, Greece, 1997, pages 126-135, 1997.
23. J. Yang, K. Karlapalem, and Q. Li, "Algorithm for materialized view design in data warehousing environment", *VLDB '97,* pages 20-40, 1997.
24. Jennifer G. Dy. 2004. Feature Selection for Unsupervised Learning, Journal of Machine Learning, pp845-889.
25. V. Rotz, and T. Lange, "Feature Selection in Clustering Problems", In Advances in Neural Information Processing Systems 16, 2003.
26. Y.Saeys, I.Inza, and P. Larr, A. N. Naga,, "A review of feature selection techniquesin bioinformatics, Bioinformatics, 23(19),, pp.2507-2517., 2007.
27. E. Baralis, S. Paraboschi, and E. Teniente, "Materialized view selection in a multidimensional database", proceedzngs of the 23rd VLDB Conference, Athens, Greece, pages 156-165, 1997.