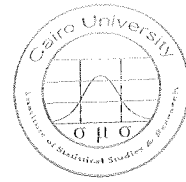




Cairo University

Institute of Statistical Studies and Research

The 42nd Annual Conference on Statistics,
Computer Sciences and Operation Research
2-5 Dec, 2007



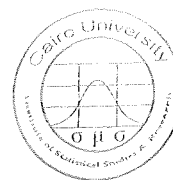
Operations Research



Cairo University

Institute of Statistical Studies and Research

**The 42nd Annual Conference on Statistics,
Computer Sciences and Operation Research
2-5 Dec,2007**





Gradient- Based Approaches for Solving Simulation Optimization Problems

Prof.Dr.\ Hegazy Zaher ¹, Prof.Dr.\Motaz Khorshid ²,
Dr.\Tawfik Younis ³ and M.Sc. \Ali Wagdy Mohamed ⁴

Abstract:

For several decades, simulation has been used as a descriptive tool by the operations research community in the modeling and analysis of a wide variety of complex real systems. With recent developments in simulation optimization and advances in computing technology; it now becomes feasible to use simulation as a prescriptive tool in decision support systems. The purpose of this paper is to provide a more comprehensive coverage of single response simulation optimization techniques with a particular focus on continuous input parameter methods, but an emphasis is placed on gradient-based approaches, discuss the major advantages and possible drawbacks of the different techniques, a case study and future research directions are also provided in this paper.

Keywords: Simulation Optimization, Single Response Simulation Optimization Techniques, Gradient-Based Approaches.

1- Introduction:

A system is a collection of entities that act and interact toward the accomplishment of a logical end. In order to study a system rigorously, the system is modeled in the form of logical and mathematical relationships. If the model is simple enough, the performance of the underlying system can be evaluated analytically.

Nevertheless, in real-world scenarios, the presence of stochastic elements and complex interactions between the system entities often preclude the possibility of obtaining an analytical solution. In these cases, the model can be studied using simulation.

Simulation is a widely used computer modeling technique that has been applied to a broad scope of problems, ranging from traffic-flow analysis to job-shop scheduling to military-campaign planning. Simulation permits the study of systems that cannot feasibly be constructed or experimented upon in the "real world," and that are too complex to be modeled analytically. When a given set of input conditions is applied to a simulation model, the model's output, referred to as a response, provides an estimate of how the true system would respond to those inputs. Although simulation is very useful in predicting the output of a system or responses, it does not in and of itself indicate the input conditions required to achieve a desired response; i.e.. it is not inherently an optimization technique, it is an evaluative methodology.

1- Department of Operations Research, Institute of Statistical Studies and Research-Cairo University.

2- Department of Decision Support, Faculty of Computer Science-Cairo University.

3- Department of Operations Research, Institute of Statistical Studies and Research-Cairo University.

4- Ph.D. Candidate in the Department of Operations Research, Institute of Statistical Studies and Research - Cairo University.

The process of finding the input conditions that yield the optimal (or near-optimal) system response(s) is referred to as simulation optimization, which can be a very expensive and time-consuming activity. In other words, simulation evaluations address "what if" questions by providing performance measures for a given set of input conditions. Whereas simulation optimization extends the evaluations by seeking optimum values for the input conditions.

The objective of simulation optimization is to determine the values of the input conditions, n controllable factors or decision variables, which optimize m responses, subject to a set of uncontrollable conditions (conditions that affect outcomes but are not under the influence of the decision maker). The simulation optimization process is complicated by the presence of random error, often the result of the combined random effects of all of the uncontrollable conditions. The response K is a random variable, whose distribution depends on the levels of the controllable factors.

Due to the presence of random error, the optimization process typically focuses on the expected value of the responses; however, while the goal is to optimize $E[y]$, only y is observable. In general, simulation optimization is in the class of stochastic optimization problems where the objective functions are stochastic functions of deterministic variables; these problems are known to be difficult to solve. In general, the responses, $Y = (Y_1, Y_2, \dots, Y_m)$, are functions of the controllable factors $X = (X_1, X_2, \dots, X_n)$, uncontrollable conditions, Z .

and random error, ϵ ; i.e., $Y = E[Y] + \epsilon = f(X/Z) = E[f(X/Z)] + \epsilon(Z)$. For example, the cost of an inventory system (Y) is a function of such controllable factors (X) as order quantity and reorder point and such uncontrollable factors (Z) as demand and lead time. Note that the additive error model considered above is only one possible model, with $E[\epsilon_i] = 0, \text{Var}[\epsilon_i] < +\infty$.

The problem under consideration is the maximization or minimization of the expected value of the objective function with respect to its constraint set as given below:

$$(\text{Max}) \min_{X \in \Theta} H(X) \quad (1)$$

Where $H(X) = E[L(X, \epsilon)]$ is the performance measure of the problem. The quantity $L(X, \epsilon)$ will be called the sample performance, ϵ represents the stochastic effects in the system, and X is a p -vector of controllable factors and Θ is the constraint set on X (feasible region).

If $H(X)$ is a one-dimensional vector, the problem is single objective (response) optimization, whereas if its dimension is more than one, the problem becomes multi-objective (multi-response). The optimum is denoted by X^* . Without loss of generality, we will consider the minimization problem throughout the paper. The above problem can be very challenging, both analytically and computationally, due to three kinds of difficulties inherent to the problem considered: lack of structure of the solution space (Θ), huge size of the solution space (Θ) and large uncertainties (ϵ) in the performance measures.

The remainder of the paper is organized as follows. Section 2 looks at gradient-based approaches, while section 3 looks at gradient estimation techniques, section 4 presents simulation metamodels. Section 5 provides response surface method (RSM) and applies it for solving simulation optimization case study. The paper ends with concluding remarks, discussion and suggestion for further research in section 6.

2- Gradient-Based Approaches:

The feasible region, Ψ , is uncountable and infinite when the set of input parameters are continuous.

An extensive body of research exists for simulation optimization problems of this type. Such continuous input parameter methods may be classified as either gradient-based or non-gradient-based approaches. In this research we will focus on gradient based approaches.

Gradient-based simulation optimization approaches have attracted a great deal of research attention over the past decade, due largely to the enormous amount of research attention given to gradient estimation techniques. Gradient-based approaches make an approximation of the objective function to be optimized by estimating the gradient ($\nabla H(X)$) of the expected system performance $H(X)$. It is assumed that the feasible region Θ is continuous. The most frequently used algorithms are Stochastic Approximation and Sample Path method.

Stochastic Approximation is designed to address the randomness arising from using simulation to evaluate the system performance and its derivatives by moving conservatively through the feasible region, so that even large errors should not put these algorithms too far off course, and they should eventually converge to the optimal solution despite the noise. On the other hand, Sample Path Optimization involves immediately converting the underlying (stochastic) simulation optimization problem to a deterministic optimization problem whose solution can be expected to be close to the solution of the original stochastic optimization problem. Then a standard deterministic (mathematical programming) optimization technique is used to solve the approximating deterministic optimization problem.

Gradient-based methods use the local gradient of the function being optimized to traverse to a locally better solution. Through iterative improvements, this procedure converges to a local optimum.

2.1 Stochastic Approximation

Stochastic Approximation (StApp) was first introduced by Robbins and Monro (1951) and Kiefer and Wolfowitz (1952). The basic assumption underlying StApp is that the original problem given by the formula $\min_{X \in \Theta} H(X)$ can be solved by $\nabla H(X) = 0$. This method uses a recursive formula: $X_{n+1} = \Pi_{\Theta}(X_n - a_n \hat{\nabla} H_n)$, where a_n is a series of real-valued step sizes that satisfy $\sum a_n < \infty$, $\sum a_n^2 < \infty$, which for example the harmonic series $a_n = a/n$ (for some constant a) satisfies. In the harmonic series sequence of step sizes, a decrease is taken at every iteration. In practice, this often leads to rather slow convergence.

The quantity X_n is the estimated value at the beginning of iteration n , and $\hat{\nabla} H_n$ is an estimate of the gradient $\nabla H(X)$ from iteration n , Π_{Θ} is a projection onto Θ . The Robbins-Monro algorithm uses an unbiased estimator (Perturbation analysis and likelihood ratios estimates) for $\nabla H(X_n)$ whereas the Kiefer-Wolfowitz algorithm uses finite-difference estimates.

The main factors are the following:

- Getting a gradient estimate $\hat{\nabla} H_n$.
- Choosing a step size a_n .
- Choosing an observation length for each iteration.

The usual assumptions needed for the convergence of the algorithm are the requirements that (1) the step size goes to zero at a rate not too fast to lead to convergence to the wrong value and not too slow to avoid convergence to a value at all, and (2) that the bias of the gradient estimate goes to zero.

As n approaches infinity, X_n approaches a value such that the theoretical regression function of the stochastic response is minimized. StApp-based algorithms are generally used in simulation optimization problems with continuous decision variables. However, SA has also been applied to discrete variable problems see for example, Gerencser (1999).

Some of the major drawbacks of this approach are its slow convergence rate, its lack of an appropriate stopping rule and its difficulty in handling constraints Shapiro (1996). Kleinman *et al.* (1999) describe the use of common random numbers for reducing the variance in stochastic approximation estimates. It has been found that, in practice, the performance of a SA-based algorithm strongly depends on the choice of the step size Fu (2002). Another disadvantage of this method is that it might find local optima, since it is based on gradient search method.

2.2 Sample Path Method

In this section, we discuss methods for continuous parameter simulation optimization that involve approximating the original simulation optimization problem (1) with a deterministic optimization problem. To illustrate, suppose that the objective function in equation (1) is of the form $H(x) = E\{z(x)\}$ (2) for all $x \in \Theta$ where $z(x)$ is a random variable having the cumulative distribution function F_x . Let N be a positive integer and assume that there exist independent random variables $Y_1, \dots, Y_N$ and a function $h: \Theta \times \mathbb{R} \rightarrow \mathbb{R}$ such that $Z_i(x) = h(x, Y_i)$ has the cumulative distribution function F_x for all $i=1, \dots, N$ and all $x \in \Theta$. Then we can approximate the objective function with $\hat{H}_N(x) = \frac{1}{N} \sum_{i=1}^N h(x, Y_i)$ for all $x \in \Theta$. (3)

Once the random variables $Y_1, \dots, Y_N$ have been generated and if each of the $h(x, Y_i)$ are independent and identically-distributed (IID) unbiased estimates of $H(x)$ then, for a sufficiently large N , the approximate objective function is $\hat{H}_N(x)$ a deterministic function of the parameter x . Therefore, we can approximate the original simulation optimization problem (1) with the deterministic optimization problem $\min_{x \in \Theta} \hat{H}_N(x)$ (4). In the simulation context, the common random numbers (CRNs) variance reduction technique provides the same sample path to calculate $\hat{H}_N(x)$ over different values of x . Now a standard mathematical programming algorithm can be applied to solve this approximate deterministic optimization problem, yielding an estimate of the optimal solution $X_N^*(\omega)$, where $\omega = (Y_1, \dots, Y_N)$.

The main advantage of the sample path optimization methodology relative to stochastic approximation is that it is capable of handling optimization problems with complicated constraints (Fu, 2002). The method also allows statistical inference and the development of stopping rules and construction of confidence intervals. Another advantage is that the method can be used with variance reduction techniques that could lead to reduction in simulation time.



2.3 Gradient Surface Method

Ho, Shi, Dai and Gong (1992) proposed the Gradient Surface Method (GSM) for optimization of discrete-event dynamic systems. GSM differs from other global search techniques because it uses traditional search methods to globally explore a response surface. It combines the advantages of RSM and efficient derivative estimation techniques such as PA and LRE in conjunction with stochastic approximation algorithms. In GSM, the gradient estimation is obtained by PA or LRE, and the performance gradient surface is fit from these gradient estimates by using least squares methods as in RSM. Zero points of the fitted gradient surface are then taken as the estimates of the optimal solution. GSM is a global search algorithm because at each iteration, it uses the information from all data points rather than just the local gradient. The most attractive features of the algorithm are that it obtains the gradient estimates by a single run, and it quickly gets to the vicinity of the optimal solution because of its global orientation. Gradient Surface Method has advantage when compared to gradient and non-gradient approaches. This method is designed for problems with multi-model response surfaces.

3- Gradient Estimation Techniques (Path Search Methods)

Naturally, the heart of gradient-based-algorithms is the technique used to estimate the gradient. Gradient estimation techniques or path search methods involve estimating a direction to move from a current feasible point that yielding an improved objective function value. Once this direction is obtained, a distance along this path is determined. The gradient-based or derivative-based methods use the first derivatives of the objective functions, or an approximation of the first derivatives, in their calculations. The gradient determines the direction of steepest ascent (descent) of the objective function at a given point in a maximization (minimization) problem.

Research in this area of simulation optimization has focused on the development of methods that are reliable as well as efficient in the estimation of gradients. Reliability is important because simulation responses are stochastic and a large error in gradient estimation may result in a movement in the wrong direction. The efficiency is a major factor because simulation experiments are expensive and it is desirable to estimate gradients with a minimum number of function evaluations.

The gradient estimation problem involves finding ways of using simulation to estimate the value of the gradient ($\nabla H(X)$) of the expected system performance $H(X)$ with respect to the input parameter x . However, gradient estimation using simulation is challenging problem that has attracted a fair amount of attention in recent years. The underlying difficulty is that if it is necessary to use simulation to evaluate the system performance, then an analytical expression for the system performance obviously does not exist and hence calculus cannot be used to obtain the required derivatives. This explains why specialized techniques are required for obtaining gradient estimates using simulation.

The two gradient based optimization approaches presented in the previous sections require estimates of the gradient of the expected system performance (simulation response) $H(x)$. There are four well-known methods in the simulation optimization literature that are used for estimating gradients of the response: (i) finite difference estimates (FD); (ii) perturbation analysis (PA); (iii) frequency domain experimentation (analysis)(FDE); and (iv) likelihood ratio estimates



3.1 Finite Difference Estimates

This method is the simplest and the most commonly used. The Finite Difference Estimate (FDE) method is based on determining partial derivatives of $H(X)$.

In order to estimate the gradient at each point, at least $p + 1$ evaluations of the simulation model are necessary. For a more reliable estimate, multiple observations for each derivative may be needed. In general, gradient estimates obtained using finite differences are used in general biased, even if the simulation results used to obtain these estimates are unbiased. This is because of error involved in approximating the gradient using finite differences. This method has the disadvantages of being computationally intensive, having slow convergence rate and large variance of the gradient estimate. The reason for this is that in order to obtain estimators with small bias it is necessary to let (ΔX) be small, but when (ΔX) is small, the estimators obtained usually have a large variance. Therefore, it is necessary to balance the desire for small bias against the desire for small variance. Several studies addressing this problem. This studies discuss how the (ΔX) should be selected to minimize the asymptotic mean-square error. These problems could be addressed by implementing common random numbers (CRN) in the different simulations to obtain a gradient estimate.

3.2 Perturbation Analysis

Perturbation Analysis (PA) was introduced by Ho *et al.* (1979) in the context of a buffer allocation problem in serial production lines. PA, when applied properly to models that satisfy certain conditions, estimates all gradients of an objective function from a *single* simulation run. There are two classifications of PA: Finite Perturbation Analysis (FPA); and Infinitesimal Perturbation Analysis (IPA). Other PA methods include IPA variants, and smoothed perturbation analysis (SPA), which is based on the "smoothing" property of conditional expectation and covers cases where IPA does not apply.

FPA is designed for discrete parameters and is a heuristic that approximates the difference in a performance measure when a discrete parameter is perturbed by one unit. IPA is used to obtain derivatives of continuous parameters and estimates all partial derivatives from a single run. In general, the idea is that in a system, if an input variable is perturbed by an infinitesimal amount, the sensitivity of the output variable to the parameter can be estimated by tracing its pattern of propagation. This will be a function of the fraction of propagations that die before having a significant effect on the response of interest. IPA assumes that an infinitesimal perturbation in an input variable does not affect the sequence of events but only makes their occurrence times slide smoothly. The fact that all derivatives can be derived from a single simulation run represents a significant advantage in terms of computational efficiency. On the other hand, the estimators derived using IPA are often biased and inconsistent.



3.3 Frequency Domain Experimentation

Frequency-Domain Experimentation (FDE) was introduced by Schruben and Coglianò (1981) for the purpose of factor screening in simulation experiments (i.e., identifying the relevant input parameters in a simulation study).

The intuitive idea behind (FDE) is to oscillate the value of the (possibly vector-valued) input parameter according to a sinusoidal function during simulation. The magnitude of the performance measure variation gives an indication of the relative sensitivity of the performance measure to the input parameter in a single simulation run, leading to estimates of the gradient of the performance measure relative to the input parameter. However, FDE has restriction that the region under investigation should be small in order to avoid large variances.

3.4 Likelihood Ratio Method

The likelihood ratio (LR) method, also known as the score function (SF). With this method the gradient is estimated by expressing the derivative of the expected value of the response with respect to an input parameter as the expected value of a function of input and simulation parameters. This value is recorded in the simulation run for a future estimation of the gradient. This method requires only one simulation run to estimate the gradient and is more generally applicable than IPA. A weakness of LR is that it may produce gradient estimates with larger variance than those obtained through IPA.

We now briefly compare the four gradient estimation techniques we have discussed. The easiest method to use is the finite-difference (FD) approach. This approach should be implemented using common random numbers (so that the variance will not be too large). The disadvantage of using this approach is that it requires several simulation runs, and it yields biased estimates of the gradient in general. Unlike finite-difference methods, both perturbation analysis (PA) and the likelihood ratio (LRE) method require knowledge of the structure of the stochastic system being simulated and hence are more difficult to use than finite-difference methods. But they require only a single simulation run and often produce estimators having desirable statistical features such as unbiased and strong consistency (unlike finite-difference approaches). The likelihood ratio method is more generally applicable than (IPA) (other variants of perturbation analysis are frequently needed in order to obtain unbiasedness or strong consistency) but often yields estimates with a large variance than the estimates obtained using IPA (assuming both techniques are applicable).

The frequency-domain experimentation (FDE) approach resembles the finite-difference approach (FD) more than it resembles the perturbation analysis and likelihood ratio approaches (except that it only requires a single simulation run). In general, it produces biased estimates of the gradient.

In general, these methods will not be of great use to us, as gradient information can be difficult to obtain from a simulation environment. A simulation model must be built and an explicit objective function must be defined. The objective function will be evaluated in terms of input parameters to the model and performance measures which are outputs from the model after execution. Obtaining gradient information involves making small changes to individual input parameters and then examining the resulting changes in the objective function.



These changes in input parameters must be made for all input parameters in order to obtain the gradient information so that the search may proceed in the most promising direction. As the search moves to a new point or region, another series of changes in the input parameters must take place and more gradient estimates must be made. This whole process becomes computationally expensive for more than just a few input parameters. And regardless of the number of input parameters, the stochastic nature of computer simulation dictates that some uncertainty will inherently exist in our estimates of the gradient at each point. This uncertainty can easily lead the search procedure astray.

Once the method of estimating the gradient is selected, one of the available search techniques can be employed to search for the optimum. One problem with using gradient-based optimization methods in optimization of stochastic simulation is that the calculated gradients are only estimates of the true gradient, because of the random error present in the observed responses. To overcome this deficiency some researchers have proposed the development of a approximations. Because approximations are models of a simulation which is itself a model of reality, they are often called metamodels or surrogate models. The goal of using a surrogate model is to provide a smooth functional relationship of acceptable fidelity to the true function with the added benefit of quick computational speed.

The approximation could be used in conjunction with a gradient-based algorithm or in entirely different ways. Once this metamodel is developed, it is used to estimate the direction of steepest ascent (for a maximization problem) or descent (for a minimization problem). The details of how to build and exploit approximations effectively keep simulation-based optimization a thriving research area.

4- Simulation Metamodels

Complex computer simulation models of proposed or existing real systems are often used to make decisions on changes to the system design. Analysts use the simulation model as a surrogate because it is impractical to construct multiple prototype versions of the real system, or because cost, time or other constraints prohibit experimentation with the real system. These models themselves may be quite complex, and so simpler approximations are often constructed; models of the model, or metamodels.

Metamodel-based optimization methods fall in two categories, both of which use an indirect-gradient optimization strategy. It is indirect because the gradient is computed for the metamodel (a deterministic function) rather than for the simulation response. Response surface methodology (RSM) is a metamodel-based optimization method that builds linear or quadratic local approximations to f to be used by a deterministic gradient based optimization strategy. Old local models are discarded and new ones are fitted at the end of each line search cycle. The global metamodel-based optimization methods build a single global metamodel (usually requiring a much larger set of simulation runs) which is then optimized using a gradient-based strategy. This section focuses on the most important metamodeling techniques:

- Response surface metamodels.
- Neural networks metamodels.
- Regression Splines metamodels.
- Spatial correlation (Kriging) metamodels.
- Radial basis function metamodels.



4.1 Response surface metamodels

In the most recent research literature, metamodeling is performed in a more localized way called sequential response surface methodology, or sequential metamodeling. Sequential RSM procedures avoid exploring the entire search space, which can be costly and often impractical. Rather, it employs linear polynomials to approximate the response surface in small sub-regions of the feasible region. In general, Response Surface Methodology (RSM) is a class of procedures that: (i) fit a series of regression models to the responses of a simulation model evaluated at several points; and (ii) optimize the resulting regression function. The basic algorithm consists of two phases. In the first phase, a first-order model is fit to the response surface. Afterwards, a steepest descent direction is estimated from the model.

The first phase is repeated until the computed slope becomes approximately zero, meaning that first-order design is no longer a good fit to the sub-region on hand. In the second phase, a quadratic response surface is fitted using second-order experimental designs, and then the optimum is determined from this fit. Canonical and ridge analysis is usually employed to examine this sub-region thoroughly with regression models.

RSM provides a general methodology for optimization via simulation. Its biggest advantage is that it uses well-known statistical tools. Compared to many gradient methods, RSM is a relatively efficient method of simulation optimization in the number of simulations experiments needed. However, It is worthy to note that even sequential RSM require a substantial amount of computational effort, particularly when the number of decision variables is large. On the other hand, RSM has a drawback due to its computational requirements if applied blindly and If the response surface of any sub-region has multiple optima, then the linear polynomial is not necessarily a good approximate. To obtain better approximate, replications with a smaller sub-region is required to provide more accurate information and after the runs are performed, there is no guarantee that the linear and/or quadratic models are going to fit the underlying response surface in an acceptable manner. This problem makes the search for optimal solutions dramatically slow and increases the computational time and without a good fit from the regression models, the gradient information will be inaccurate and the search will proceed an incorrect direction.

4.2 Neural networks metamodels

Neural networks can be thought of as flexible parallel computing devices for producing responses that are complex functions of multivariate input information. They can approximate arbitrary smooth functions and can be fitted using noisy response values. Neural networks are networks of numerical processors, whose inputs and outputs are linked according to specific topologies. Networks used for function approximation are typically multi-layer feed-forward networks. Feed-forward layered networks have the flexibility to approximate smooth functions arbitrarily well, provided sufficient nodes and layers.

A simulation model can be used to generate responses associated to a specific configuration of input parameters and then utilize them to train the network. In this way it is possible to create a "black box" representing an approximated function of the simulation model. This black box can be used integrated with a search heuristic to find better responses and eventually the simulation optimum.



4.3 Regression Splines metamodels

Splines, where the polynomials are defined in a piecewise manner rather than as a single expression for the entire data set, provide an alternative solution to the problem of fitting highly nonlinear data. Simply put, several low order polynomials are fit to the data, each over a separate range defined by the knots or break points. Then, boundary conditions are placed on the polynomials to ensure that the pieces match up with a prescribed order of continuity. In general, if a linear or quadratic polynomial regression does not provide a good fit, it is natural to think of higher-order polynomial approximations.

4.4 Spatial correlation (Kriging) metamodels

Sacks et al. (1989) and numerous references therein develop a spatial correlation parametric regression modeling approach. The expected smoothness of the function is captured by a spatial correlation function. Spatial correlation models, also called kriging models, have recently become popular for deterministic simulation metamodels. They are more flexible than polynomial models in fitting arbitrary smooth response functions, and seem to be less sensitive than radial basis functions to small changes in the experiment design.

4.5 Radial basis function metamodels

Radial basis functions (RBF) provide an alternative approach to multivariate metamodeling. In an empirical comparison, Franke (1982) found radial basis functions to be superior to thin plate splines, cubic splines, B-splines, and several others for fitting deterministic response functions.

The radial basis function approximation consists of a sum of radially symmetric functions centered at different points in the domain Θ .

Metamodel-based optimization simplifies dealing with issues that complicate direct optimization of the simulation model, such as multiple local optima, multiple objectives, and constraints on design parameters and/or responses. This is because the implicitly represented stochastic response of the simulation is replaced by an explicit deterministic metamodel response function.

A metamodel-based optimization strategy consists of choosing a metamodel form, designing an experiment to fit the metamodel, fitting the metamodel and validating the quality of its fit, optimizing the metamodel (or using it to provide a search direction), and checking the performance of the simulation at the metamodel-predicted optimum (or in the metamodel-determined search direction). In some cases this process is repeated, with the new experiment design focused on the neighborhood of the predicted optimum. Two general strategies have been used for metamodel-based simulation optimization: global metamodel fit, followed by optimization, and iterated local metamodels.

The iterative local metamodeling strategy is commonly used to determine an optimization search direction. This is followed by a line search using the simulation response directly. Because the metamodels are local, Taylor's Theorem implies that linear and quadratic polynomial models can provide adequate fit. This is the scenario for response surface methodology. Of course, determining the meaning of 'local' is critical to the adequacy of these metamodels and to the success of the method. If the local region is too small, response surface characteristics will be swamped by variation in the simulation model output. If the local region is chosen too large, linear or quadratic approximations will be inadequate.

Global metamodel fits using polynomial response surface metamodels are rarely adequate. Instead, spline, neural network, spatial correlation or radial basis function approximation is recommended. The experiment designs for fitting global approximation metamodels differ from the central composite, factorial and fractional factorial designs of RSM optimization. Orthogonal array-based Latin hypercubes tend to perform well. Since global metamodels can have multiple local optima, a global optimization strategy is recommended.

5 Case Study

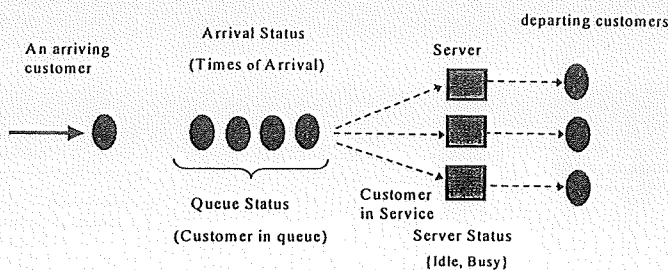
To validate the Response Surface Method, a simulation model of Multiple server models $\{(M/M/c):(FCFS/\infty/\infty)\}$ queuing system has been developed by (C++ programming language) and used with the objective of finding the appropriate model which approximate the relation between the expected average waiting time in queue (w_q) as the response(dependant variable) and service rate and number of servers as input factors (independent variables), then finding the optimal service rate μ and the optimal number of servers C that minimize the expected average waiting time in queue (w_q).

5.1 Problem analysis:

Simulation of a Multiple Server System

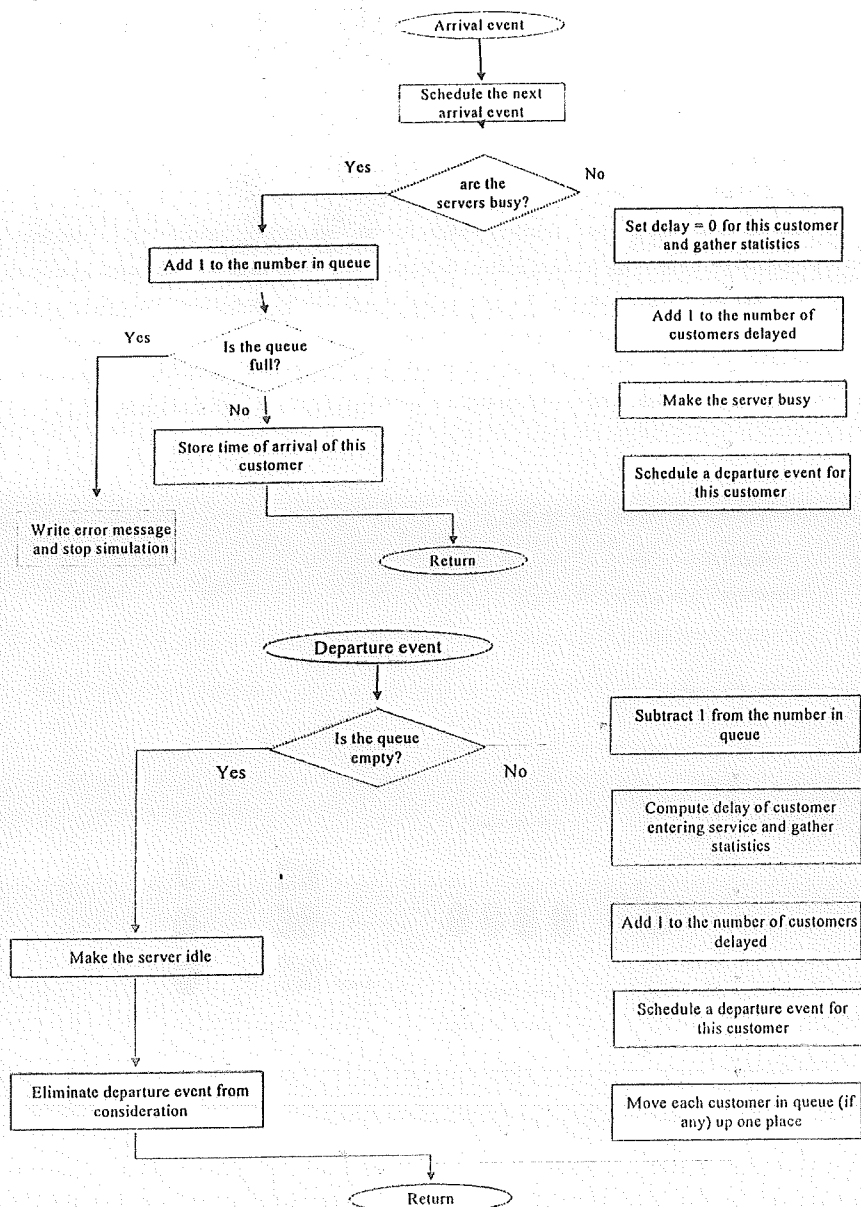
- Problem Statement
- Program Organization and Logic
- Problem formulation

Problem Statement



(Simulation modeling and analysis averill, kelton, 2002)

Program organization and logic



Problem formulation

The stochastic output (response) is:

Expected average waiting time in queue (w_q).

The input factors are:

- The service rate (μ).
- The number of servers (N).

The objective is:

finding the appropriate model which approximate the relation between the expected average waiting time in queue (w_q) as the response (dependant variable) and service rate and number of servers as input factors (independent variables), then finding the optimal service rate μ and the optimal number of servers C that minimize the expected average waiting time in queue (w_q).

5.2 Solution steps and results:

Apply proposed framework to simulation model and using software packages such as (Minitab, Jmp, and Winqsb) to analyze and optimize the model.

Step (1):

It is concluded that the Expected average waiting time in queue (w_q) (AWT, in minute) of queuing process is main affected by the service rate (x_1 , in minute) and by the number of servers (x_2 , number). The region of operation is $1 \leq x_1 \leq 5$, $1 \leq x_2 \leq 5$.

Factor	Low	Center	High
X1	1	3	5
X2	1	3	5

Now, approximate the simulation response function (AWT) in the current region of interest by first order-model.

Decision: Based on the analysis results.

The regression equation is

$AWT = 0.856 - 0.121 X_1 - 0.127 X_2$, $Mse = 0.2504$ and $R-Sq (adj) = 42.1\%$. The P-value for the overall lack of test is ($P=0.00$), implying that the linear model is not adequate.

There is evidence that a linear regression model is not appropriate, which implies that the optimization should proceed with Phase II that is, the fit and optimization of a second -order model.



Step (2):

Since the relation appearing is not linear, approximate the simulation response function(AWT) in the current region of interest by second-order model.

Decision: Based on the analysis results.

The regression equation is

$$AWT = 29.8 - 23.3 X_1 + 8.4 X_2 + 0.06 (X_1 * X_2) + 2.94 (X_1)^2 - 1.31 (X_2)^2.$$

Mse= 13.76, R-Sq (adj) = 34.3%, The P-value for the overall lack of test is(P=0.00),implying that the quadratic model is not adequate.

Step (3):

Try possible transformations to response and variables. It is clear that the inverse transformation to independent variables is good (since the relation between the response and independent variables is inverse relation).

Decision: Based on the analysis results.

The regression equation is

$$Aw_t = 0.344 - 0.872(1/X_1) - 1.04(1/X_2) + 0.324(1/X_1)^2 + 0.514(1/X_2)^2 + 1.63(1/X_1 * X_2).$$

Mse=0.06144, R-Sq (adj) = 90.6%, The P-value for the overall lack of test is(P=0.00),implying that the model is not adequate.

There is evidence that the model is not appropriate.(must try a good transformation for response).Now consider the same data with $\log_{10}$ (AWT).

Step (4):

Repeat the experiment with $\log_{10}$ (AWT), $\log_{10}$ (response).

Decision: Based on the analysis results.

The regression equation is

$$\text{LOG (AWT)} = - 9.29 + 2.98 (1/X_1) + 17.6 (1/X_2) - 0.25 (1/X_1)^2 - 10.2 (1/X_2)^2 - 0.918 (1/X_1 * X_2),$$

Mse=0.3226, R-Sq (adj) = 96.3%, No evidence of lack of fit ($P > 0.1$).

The model is adequate. But eliminate the non significant terms from the model and repeat the analysis.

The graphs give a definite signal that curvature is present with $1/x_2$ and the term $(1/x_2)^2$ is significance.using stepwise and best subset method to testing a subset of the regression coefficients.

Decision:based on the analysis, the best model which contains the intercept,

$1/X_1$, $1/X_2$ and $1/(X_2)^2$ terms.since if the term $1/X_1 * X_2$ included will not improve the model statistics sufficiently.

Step (5):

After elimination the nonsignificant terms from the model.

Decision: Based on the analysis results.

The regression equation is

$$\text{LOG (AWT)} = - 9.03 + 2.25 (1/X_1) + 17.2 (1/X_2) - 10.2 (1/X_2)^2,$$

Mse=0.3186, R-Sq (adj) = 96.3%, No evidence of lack of fit ($P > 0.1$). The model is adequate.



Step (6):

Optimize the process using winQSB. (Estimated solution)

Experimental results

MINIMIZE - 9.03 + 2.25 (1/X₁) + 17.2 (1/X₂) - 10.2 (1/X₂)²

X₁ >=1, <=5 X₂ >=1, <=5

Solution Summary for Queuing Model Optimization

Decision Variable	Solution Value
-------------------	----------------

1 X ₁	4.9962
----------------------	--------

2 X ₂	4.9985
----------------------	--------

Minimized Objective Function = -5.5469

So X₁ = 5, X₂ = 5 approximately. Log₁₀ (awt) = -5.5469, awt (wq) = 0.000003 min.

Step (7):

Solve the process using winQSB. (Analytical solution)

Queuing model analysis

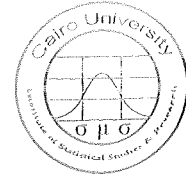
Arrival rate = 0.5 customer per min.

Service rate = 5 customer per min.

No of servers = 5.

No of customer = 1000000.

(w_q) = 0.0000 min.



6. Discussion and further research directions

In the past, simulation has often been regarded as the method of last resort among OR researchers and users. However, over the years as a result of advances in computer technology and simulation languages, it has become one of the most valuable OR tools (it is usually ranked at the top of the list). The need for dynamic and stochastic models to analyze complex systems is also expected to increase the interest for simulation in the future. A simulation model is typically a descriptive model of the system, i.e., it describes the behavior of the system under consideration, and helps us to understand the dynamics and complex interactions among the elements of the system. In the scientific community, simulation has often been criticized for a lack of optimization capability; simulation results have typically been runs with a set of observations, rather an optimum solution to the problem as usually found by the prescriptive or normative models (e.g., linear programming, dynamic programming). By incorporating optimization features in simulation systems, we not only eliminate its major limitation, but also open a door for simulation in terms of new application areas and research possibilities. With the ability of performing optimization, simulation can also become an operational tool to solve short-term decision-making problems as well as strategic and tactical problems (i.e., design and long-term planning problems).

In this paper, we presented a comprehensive coverage of single response simulation optimization techniques with a particular focus on continuous input parameter methods, but an emphasis is placed on gradient-based approaches. The effectiveness of (RSM) is demonstrated by applying it to a simulation model of multiple server queuing system which has analytical solution.

- a) By comparing the analytical and estimated solutions, it was found that there is no difference between them. This would provide empirical evidence that the optimization method is relatively robust.
- b) The response surface method can be used to optimize a different type of queuing systems and other simulation models which has no analytical solution.

The methods proposed for simulation optimization each carry their own assumptions, and the performances depend on the characteristics of the problems to which they are applied.

In the literature, there have been also some research efforts to compare the methods. The conclusions from these studies vary, i.e., one method can perform very well for one problem but can be the worst for another problem. In our opinion, one future research effort would be to generate benchmark problems or define some standard test functions to evaluate performances of different techniques. These test problems also provide important insights into which methods to integrate for solving a particular problem.

As discussed in Law and McComas (2000), some commercial simulation companies have already integrated their simulation systems with some kind of global search methods (SA, GA, and tabu search). In this context, these benchmark problems can also serve as a platform to evaluate the effectiveness (solution quality and run times) of future simulation packages.



Since the problems that are to be solved by simulation optimization vary in terms of the number and structure (i.e., discrete or continuous, quantitative or qualitative) of decision variables, and shape of the response function, there is no single method to solve all of these problems. This forces researchers to develop more robust techniques that can handle a larger class of problems. For instance, EAs require no prior idea about the response surface and can handle both quantitative and qualitative variables. Therefore, they seem to be very promising in terms of future research potential in this area.

As pointed out by Andradóttir in Fu *et al.* (2000), simulation optimization still requires a considerable amount of computer time. This is one of the important barriers on future simulation optimization applications since a model has to be run under a large number of experimental conditions. One possible solution (or at least away to alleviate the problem) would be to use parallel processors or the utilities of distributed computing. There are already some commercial simulation products that make use of the benefits of distributed computing in a multiprocessing environment.

The other possibility to reduce the computational burden, as suggested by Harrell in Fu *et al.* (2000), would be to use statistical methods such as factor screening techniques to reduce the search space. More research also needs to be conducted in these areas to improve the efficiency of simulation optimization systems. It appears that most of the existing simulation software packages with optimization features rely on meta-heuristics (SA, GA, tabu, etc.). However, there is already a great deal of work that has been done in RSM and gradient-based methods. In future studies, these traditional methods together with meta-heuristics can be jointly used to develop better hybrid approaches in dealing with simulation optimization problems.



References

- 1-Franke, Richard 1982. Scattered data interpolation: tests of some methods. *Mathematics of Computation*, 38, 181-200.
- 2-Fu, M.C., Carson, J.S., Harrel, C.R., Kelly, J.P., Andrad' otir, S., Glover, F.G.,Ho, Y. andRobinson, S.M. (2000) Integrating optimization and simulation: research and practice, in *Proceedings of the 2000 Winter Simulation Conference*, IEEE Press, Piscataway, NJ, pp. 610-616.
- 3-Fu, M. (2002). "Optimization for Simulation: Theory and Practice," *INFORMS Journal on Computing*, Vol. 14. No. 3, pp. 192-215.
- 4-Gerences'er, L. (1999) "Rate of convergence of moments for a simultaneous perturbation stochastic approximation method for function minimization." *IEEE Trans. Automatic Contrl.*, Vol. 44, pp. 894-906.
- 5-Ho, Y.C., Eyler, A. and Chien, T.T. (1979) A gradient technique for general buffer-storage design in a serial production line. *International Journal of Production Research*, 17, 557-580.
- 6-Ho, Y.C., Shi, L., Dai, L. and Gong, W.B. (1992) Optimizing discrete event dynamic systems via the gradient surface method.*Discrete Event Dynamic Systems*, 2, 99-120.
- 7-Kiefer, J. andWolfowitz, J. (1952) stochastic estimation of the maximum of a regression function. *Annals of Mathematical Statistics*, 23, 462- 466.
- 8-Kleinman, N.L., Spall, J.C. and Naiman, D.Q. (1999) Simulation-based optimization with stochastic approximation using common random numbers. *Management Science*, 45, 1570-1578.
- 8-Kiefer, J. andWolfowitz, J. (1952) stochastic estimation of the maximum of a regression function. *Annals of Mathematical Statistics*, 23, 462- 466.
- 9-Law, A.M. and McComas, M. G. (2000) Simulation-based optimization, in *Proceedings of the 2000 Winter Simulation Conference*, IEEE Press, Piscataway, NJ, pp. 46-49.
- 10-Robbins, H. and Monro, S. (1951) A stochastic approximation method. *Annals of Mathematical Statistics*, 22, 400-407.
- 11-Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P. (1989). Design and analysis of computer experiments. *Statistical Science* 4, 409-435.
- 12-Schruben, L.W. and Cogliano, V.J. (1981) Simulation sensitivity analysis: a frequency domain approach, in *Proceedings of the 1981 Winter Simulation Conference*, IEEE Press, Piscataway, NJ, pp. 455-459.
- 13-Shapiro, A. (1996). Simulation based optimization. *Proceedings of the 1996 Winter Simulation Conference*, pp. 332-336.

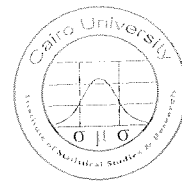


Cairo University

Institute of Statistical Studies and Research

The 42nd Annual Conference on Statistics,
Computer Sciences and Operation Research
2-5 Dec, 2007





Using Artificial Neural Networks to Minimize Errors in Distance Recognition

Hegazy Zaher *

Mohammed Ewees**

Taher Taha Zaki***

Abstract

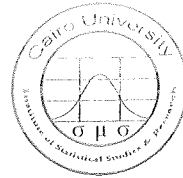
We describe a neural approach to determine the distance of vehicle based on sound emitted by them .The sound is preprocessed using wavelet method and Fourier transform method to obtain feature vectors to be used by neural classifier. We use Backpropagation, self organizing neural network for classification. By doing simulation for neural networks we find feature vectors extracted from wave let method that input to ANN achieves high accuracy then vector extracted from Fourier transform, after design both Backpropagation and SON and test both of them we find Backpropagation high accuracy then SON in our case, due to change of propagation of wave at day morning and night due to noise we record the sound produced by vehicles using microphone to determine the distance at different time of day. We record 1600 samples for vehicle sound using microphone at distances of 20 cm and 1 meter 2meter and 3 meter (400 for each distance position) to design the ANN .

Key wards: Artificial neural networks, Distance recognition, Minimization of errors.

* Professor of Math. Statistics, Cairo University

**Assistant professor of Operations Research, Cairo University

*** Ms Student of Operations Research, Cairo University



1-Introduction

A sound generated by vehicle is basis for classification process and consider as source of signal. The recognition divided into two stages: preprocessing to extract the specific features vector from signal and classification to compare the feature vector with master vectors defined during learning procedure and to make final decision[1]. The signals converted to computer files using microphone and basis to construct the masters vectors related to each position of the distances that vectors are the input to ANN, then speed, distance, type of vehicle, and extra parameters can be determined based on sound emitted by vehicles.

2-The Features Extracting

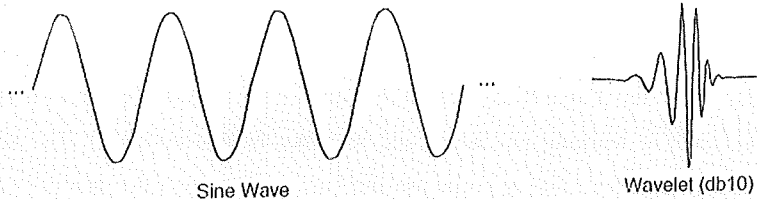
2.1- Fourier Transform

Fourier analysis breaks down a signal into constituent sinusoids of different frequencies, can be use for preconditioning data for training neural networks .From practical viewpoint the fast (FFT) temporal signals is an efficient method to carry out discrete Fourier transform DFT the FFT is a built in function in Matlab A major advantage of using the magnitude Spectrum of the FFT that is insensitive to the of signal .Therefore if the phase of the signal is not important, the FFT magnitude samples can be used as feature vector for each training pattern (signal). Another way to think of Fourier analysis is as a mathematical technique for transforming our view of the signal from time based to frequency based [2].

2.2- Wavelets and wavelet Transforms

What is a wavelet?

A wavelet is a waveform of effectively limited duration that has an average value of zero as in fig (1) [3].



Fourier analysis is used for analyzing waves that certain functions or signals can be expanded in terms of sinusoidal waves .This method of signal analysis is very useful when we deal with physical data that are periodic time invariant or stationary . A wave let can provide a useful tool for analyzing signals that are time varying, transient or no stationary. Thus many type of signals that can not be properly analyzed using Fourier methods can be investigated using wavelet approach. The wavelet transform can provide a time-frequency description of signals and can be used to compress data for training neural networks[2].

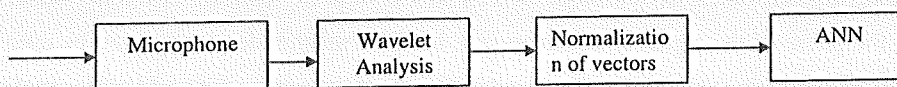


Fig (1) System for signal processing based on wavelet

We use both Fourier analysis and Wavelet analysis and we find the wavelet method achieves high accuracy then Fourier method.



2.3-Signal preprocessing

We record the sound emitted by vehicle by microphone then we use both discrete wavelet transform DWT and Fourier Transform to extract feature vectors using mat lab program. We do not use continuous wave transform because its coefficient generates awful lot of data.

To record wave (sound) produced by vehicles using the following matlab commands:

```
>> fs = 11025;  
>> wav201 = wavrecord (1*fs,fs,'double')  
>> save wav201
```

where fs Standard audio rates are 8000, 11025, 22050, and 44100 Hz.

2.3.1 Matlab code to extract feature vector using Fourier analysis

For each distance we write the following code to extract feature vectors

```
>> x201 = fft(wav201,11);  
>> px201 = x201.*conj(x201)    % use power spectrum to eliminate complex number
```

2.3.2 Discrete wavelet transform

For many signals, the low-frequency content is the most important part. It is What gives the signal its identity? The high-frequency content, on the other hand, imparts flavor or nuance. Consider the human voice. If you remove the high-frequency components, the voice sounds different, but you can still tell what's being said. However, if you remove enough of the low-frequency components, you hear gibberish. In wavelet analysis, we often speak of approximations (cA) and details(cD). The approximations are the high-scale, low-frequency components of the signal. The details are the low-scale, high-frequency components. In wavelet packet analysis, the details as well as the

approximations can be split. This yields 2^{n-1} more than different ways to encode the signal. This is the wavelet packet decomposition tree as in (fig 2).

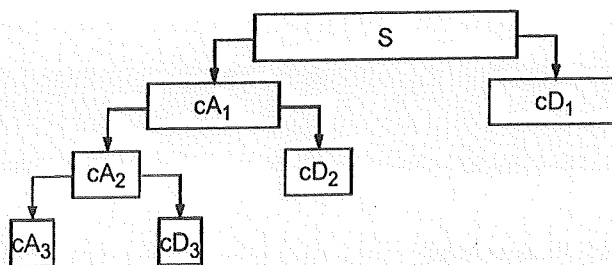


Fig (2)The wavelet decomposition tree

2.3.3 Matlab code to extract feature vectors using wavlet

For each distance we write the following code to extract feature vectors

1-load wave wav20

```
s = wav20(1:11025);
```

```
l_s = length(s);
```

2-Perform a multilevel wavelet decomposition of a signal.

To perform a level 10 decomposition of the signal (again using the db1 wavelet), type

```
[C,L] == wavedec(s,10,'db1');
```

3- Extract approximation and detail coefficients.

cA11 = appcoef(C,L,'db1',10'); that are 11 rows represent input vector to ANN for each distance.

To extract the levels 3, 2, and 1 detail coefficients from C , type

```
[cD1,cD2,cD3] == detcoef(C,L,[1,2,3]);
```

3. Neural Networks

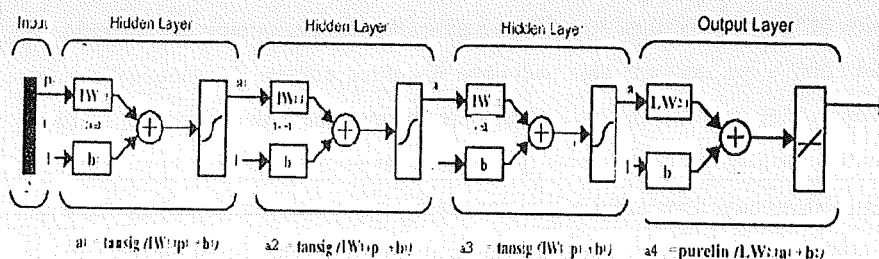
We propose two different neural networks for solving distance recognition. The first is backpropagation neural network. The second approach is self organizing neural network.

3.1 Backpropagation Neural Networks

In This section ,we introduce the architecture of the network that is most commonly used with the backpropagation algorithm , the multilayer feedforward network.

The network receives the input vector contain 11 rows and one column to identify the distance for vehicle by responding with 1 element output.

The neural network needs 11 inputs and 1 neuron in output layer with linear transfer function to identify the distance. The network is three layer tan-sigmoid network in hidden layers because its output in range (-2 to 2). The first hidden layer 24 neurons, second is 14 and third is 6 neurons . This number picked by guesswork and experience. We record 1600 sound produced by vehicle at distances 20 cm , 1 meter, 2 meter and 3 meters each distance have 11 vectors represent low frequency of signals produced by vehicles (approximations). First step we load the data into workspace and perform principal component analysis. to remove redundancy in data ,next step to divide the data into training validation and test subsets.





3.1.2 Recognizer

A network is to be designed and trained to recognize the matrix of the 4 distance position, the matrix contain 11 rows and 400 columns for each distance. i.e. the matrix contain 11 rows and 1600 columns. Each column represent vector for specified distance.

3.1.3 Preprocessing and Post processing

Before training we want to scale network inputs and targets to normalize the mean and standard deviation of the training set. Writing the matlab program code

```
[pn,meanp,stdp,tn,meant,stdt] ==prestd(p,t);
```

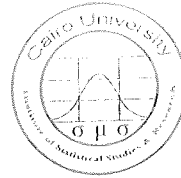
If prestd is used to scale both the inputs and targets, then the output of the network is trained to produce outputs with zero mean and unity standard deviation. If you want to convert these outputs back into the same units that were used for the original targets, then you should use the routine poststd . In the following code we simulate the network that was trained in the previous code, and then convert the network output back into the original units[4].

```
an =sim(net,pn);
```

```
a =poststd(an,meant,stdt);
```

3.1.4 Principal Component Analysis (prepca, trapca)

In some situations, the dimension of the input vector is large, but the components of the vectors are highly correlated (redundant). It is useful in this situation to reduce the dimension of the input vectors. An effective procedure for performing this operation is principal component analysis. This technique has three effects: it orthogonalizes the components of the input vectors (so that they are uncorrelated with each other); it orders the resulting orthogonal components (principal components) so that those with the largest variation come first; and it eliminates those components that contribute the least to the variation in the data set. The following code illustrates the use of prepca , which performs a principal component analysis [4].



```
[pn,meanp,stdp]==prestd(p);  
[ptrans,transMat]==prepca(pn,0.02);
```

3.1.5 Post-Training Analysis (postreg)

The performance of a trained network can be measured to some extent by the errors on the training, validation and test sets, but it is often useful to investigate the network response in more detail. One option is to perform a regression analysis between the network response and the corresponding targets. The routine postreg is designed to perform this analysis. The following commands illustrate how we can perform a regression analysis on the network that we previously trained in the early stopping section[4].

```
a=sim(net,p);  
[m,b,r]=postreg(a,t).
```

3.1.6 Matlab code to design back propagation

```
Load vmat; % file contain data  
[pn,meanp,stdp,tn,meant,stdt]=prestd(p,t);  
[ptrans,transmat]=prepca(pn,0.00001);  
[r,q]=size(ptrans)  
iitst=2:4:q; iiival=4:4:q;  
iitr=[1:4:q 3:4:q];  
val.P=ptrans(:,iiival);  
val.T=tn(:,iiival);  
test.P=ptrans(:,iitst);test.T=tn(:,iitst);  
ptr=ptrans(:,iitr); ttr=tn(:,iitr);  
net=newff(minmax(ptr),[24 14 6 1],{'tansig' 'tansig' 'tansig' 'purelin'},'trainlm');  
[net,tr]=train(net,ptr,ttr,[],[],val,test);
```

```

legend('Training','Validation','Test',-1);
ylabel('Squared Error');xlabel('Epoch')
an=sim(net,ptrans);a=poststd(an,meant,stdt);
for i=1:1; figure(i)
[m(i),b(i),r(i)]=postreg(a(i,:),t(i,:));
end

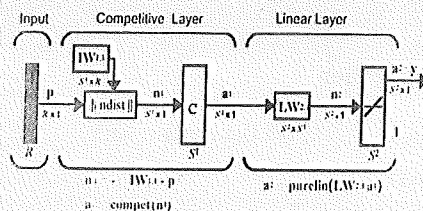
```

When we run the program the accuracy of stored data is 90% percent

When we run the program online the accuracy of online data is 70 %

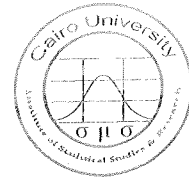
The difference between tow result due to noise which effect on the propagation of sound wave.

3.2 Learning Vector Quantization Networks



It has been postulated that brain uses spatial mapping to model complex data structures internally. Kohonen uses this idea to good advantage in his network because it allows him to perform data compression on the vectors to be stored in network, using a technique known as vector quantization [5].

LVQ network has first competitive layer and second linear layer the competitive layer learns to classify input vector as follows the input vector and input weight matrix produces vector having s elements. the elements are the negative of distances between the input vectors IW formed from the rows of input weight matrix. the net n of competitive layer is computed by finding the negative distance between input vector p



and weight vectors adding the biases b . if all biases are zero ,the maximum net input a neuron can have is 0. This occurs when p equal neurons weight vector. The competitive transfer functions accepts a net input vector for a layer and returns neuron outputs of 0 for all neurons except the winner the neuron associated with most positive elements of net input n the winner's output is 1. if all biases are 0, then the neuron whose weight vector is closest to input vector has at least negative net input and therefore wins the competition to output a 1. The linear layer transforms the competitive classes into target classification defined by user. We refer to the classes learned by competitive layer as subclasses and the classes of the linear layer as target classes [4].

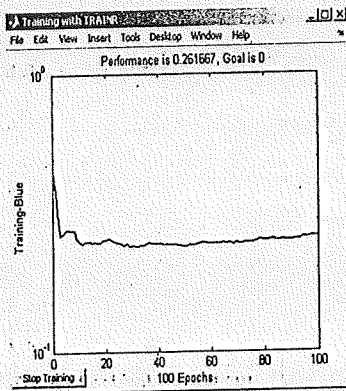
3.2.1 Matlab code to design SON

```
p=[m20 m80 m200 m300];  
t1=ones(1,400)  
t2=2*ones(1,400)  
t3=3*ones(1,400)  
t3=4*ones(1,400)  
tc=[t1 t2 t3 t4]  
t=ind2vec(tc)  
targets=full(t)  
net=newlvq(minmax(p),10,[0.25 0.25 0.25 0.25],0.01)  
net.trainparam.epochs=100  
net= train(net,p,t)  
y=sim(net,p)  
ycc=vec2ind(y)
```

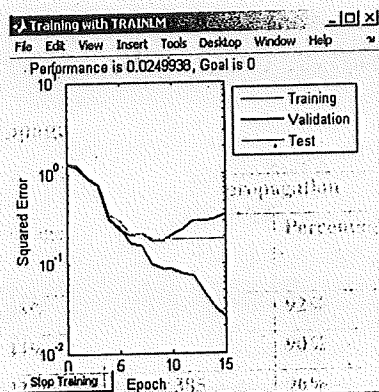
4-Experiments

4.1 Comparison between SON and Back propagation networks

Patterns	distance	Using SON		Using Back propagation	
		Number of correct	Percentage	Number of correct	Percentage
From 1: 400	20 cm	194	63%	370	92%
From 401:800	100 cm	153	44%	360	90%
From 801:1200	200 cm	400	22%	385	96%
From 1201:1600	300 cm	400	83%	370	92%



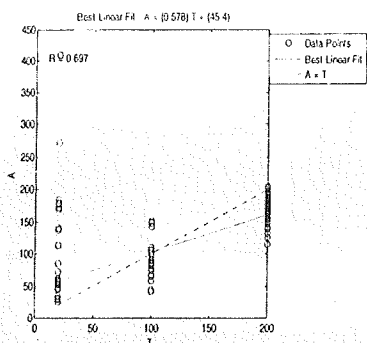
Training with SON



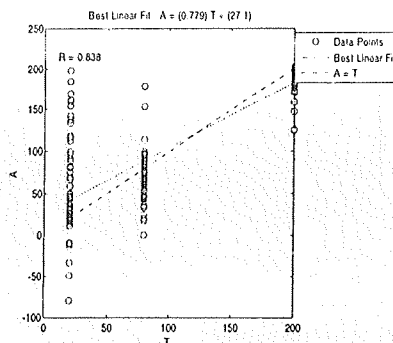
Training with Back propagation

4.2 Comparison between FFT and wavelet method

To compare between FFT and Wave let methods we extract 225 feature vectors using both methods and use them as input to same ANN .then measure the accuracy (the difference between actual output and targets(desired output)) of ANN.



Accuracy when using FFT(70%)



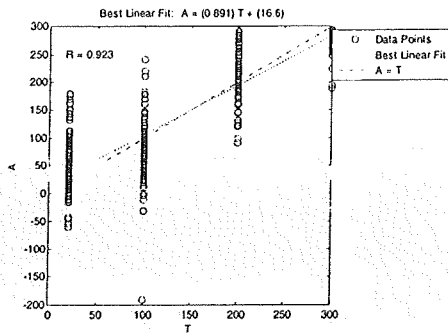
Accuracy when using wavelet (84%)

We find the Accuracy when using wavelet is higher then Accuracy when using wavelet.

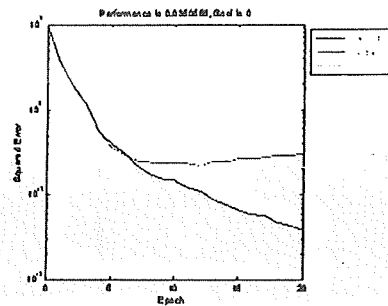
4.3 Experiment using Backpropagation

The ANN is trained using 800 patterns and tested with 400 patterns and made validation with 400 patterns.

Patterns	distance	Using Back propagation	
		Number of correct	Percentage
From 1: 400	20 cm	380	95%
From 401:800	100 cm	340	85%
From 801:1200	200 cm	375	93%
From 1201:1600	300 cm	380	95%
Total percentage			90%



Accuracy of neural network
(92% percentage)



Training with backpropagation

5-Conclusion

This paper demonstrates how distance recognition can be designed Using artificial neural networks to minimize error in distance recognition. The achieved results are better for backpropagation then self organizing network in our case. The other area of future work is the problem of selecting number of hidden neurons in both networks.

References:

- 1- Henry Maciejewski, Jacek Mazurkiewicz ,krzysztof Skowron, Tomaz Walkowiak , Neural Networks for Vehicle Recognition, institute of Engineering Cybernetics, ul.janiszewskiego 11/17, 50-372 Wroclaw, POLAND e-mail :twalkow@ictpwr.wroc.pl.
- 2- Fredric M. Ham, Ivica Kostanic, Principle of NeuroComputing for Science and Engineering.,Mcgraw hill ,2001.
- 3- Michel Misiti,Yves Misiti,Georges Oppenheim,Jean-Michel Poggi,Wavelet Toolbox For Use with Matlab, copyright 1997 - 2001 By The MathWorks, Inc)
- 4- Howard Demuth,Mark Beale, Neural Network Toolbox For Use with Matlab, copyright 1997 - 2001 By The MathWorks, Inc)
- 5- R Beale and T Jackson, Neural Computing: An Introduction, Adam Hilger, copyright 2000.



Designing Survivable Fiber-Optic Networks

With Application to the Egyptian Backbone.

Taghrid Abdel Rahman Imam¹

Abstract

During the last years fiber optic based communication networks have become a fordable for individual households. Due to the costly nature of the materials and infrastructure that have to be deployed to distribute such a network over a large area, exact algorithms for calculating a provably optimal cable routing are desirable. However, such algorithms are not always able to determine an optimal solution within acceptable time. Heuristics and meta heuristics can provide techniques to calculate approximate solutions that are often sufficient as a starting point for certain, more time consuming approaches, or even sufficient for practical purposes.

Networks survivability is the most concerning issue in today's optical network, because any failure will cause a great loss of information, so it will affect the Quality of Service. In this paper network survivability models are built, where Mean Time Between Failures (MTBF) and Mean Time To Repair (MTTR) are measured, depending upon the components types of the fiber optics stations, and the failure rates of fiber optics cables. Mesh architecture is proposed, and analyzed to increase the survivability of the national networks such as the Egyptian backbone when applying the span protection, and path protection method on this architecture.

(Key Points: Network modeling, Graph theory, Networks reliability and availability, and Terms-Protection technique).

1. Introduction:

High survivability of the telecommunication services has traditionally been expected and is being equally important for both end users and contracted services providers. The importance of network survivability grows with steadily increase of the network capacity. Next generation optical networks will provide the pathways to transport, potentially, terabits of data.

¹ Al Azhar University, Faculty of Science for Girls

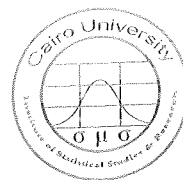


Network reliability and protection are among the most important issues concerning the transport of high-speed connections: interruption of an optical connection even for a short period could cause the loss of a huge quantity of data (e.g. 5 GByte for 40 Gbit/s modulated wavelength channel) [6, M. Tornatore, G. Maier, A. Pattavina, 2005], failures happen quite frequently and with catastrophic consequences. The impact of the network outage can be normally measured in terms of customer-minutes. Outage can be normally multiplied by the number of affected customers.

In today's telecommunication industry much effort is put into the increase of infrastructure. This paper studies the network survivability optimization and gives guidance to network operators on how to increase the quality of service. To survive today's marketplace, operators must run efficient, network survivability optimization to make investment and repair strategy with respect to their Service Level of Agreement, (SLA). As part of SLA the network operator commits certain survivability for a connection. Common requirements are that a connection be available 99% to 99.999% (five 9s) of the time.

Five 9s corresponds to a connection downtime of 5 minutes per year. In order to achieve high service survivability networks are designed survivable, they are able to continue providing service in the case of failure. The involved techniques of protection and restoration are explained in section 2. They use a pre-assigned or dynamically assigned spare capacity within the network enabling the rerouting of the affected traffic around the failure. Networks are commonly designed for 100% protection or restoration if a single failure is occurring to any span or path, this can be handled but node failures are not considered. This may lead to 100% protection or restoration for a single failure but not for multiple failures [2, G. Birkin, J. Kennington, E. Olinick, A. Ortyński, and G. Spider, 2005].

Failures in an optical network can be distinguished whether they damage links or switching devices. In the first situation, faults often results from external causes: cable cuts are very frequent especially in the terrestrial networks since fiber cables often share other utility transport conduits, such as gas or water pipes and electrical cables. Equipment failures in the network nodes are mainly due to internal causes, such as hardware degradation or management software inefficiency. They can result also from exceptional events such as natural phenomena, power block outs. As in the report of the U.S. during 1997, there are 136 cuts per year, so, it is irrelevant in some places to take the optical cuts only. The equipments are taken into consideration. Some authors mentioned that equipment failures are proved to be less common on average than transmission links failures.



In our study we will compare between links failures and equipments failure. [2, G. Birkin, J. Kennington, E. Olininck, A. Orzynski, and G. Spider, 2005]

Survivability calculations demonstrate which area needs to be protected according to the digging-up, also according to the topology of the network, whether they are ring or mesh, because for many years, ring and mesh architecture have been under consideration by operators for their networks. In the study of network architecture vulnerability (availability), it was found that mesh architecture is more robust than two rings connected to each other, for a single ring and star whatever is the number of nodes [1, Dalia M. Elgamel, Prof Dr. Atalla Hashad, 2007], [10, W.D. Gover, 2003]. Rings can restore failures fast in (50-60ms), with at least 100% extra capacity where this figure is only 50-70% (depending on the network topology) for meshes. In this paper, we deal with Optical Transport Networks (OTNs), survivability calculations using span protection, & path protection technique.

The software is used to enhance the survivability of a network according to its topology, using graph theory algorithms. The survivability which we refer to is the classical reliability theory [8, S. Ramamurthy, L. Sahasrabudhe, Biswanath Mukherjee, 2003]; also, the network design based on static traffic matrix.

2. System Reliability and Survivability:

The reliability and survivability analyzes the system based on a set of distinct subsystems, connected to obtain an intended function. Reliability (R) is defined as the probability that a system will perform its intended function for a specified period of time under a given set of conditions. Availability (A) inherently reflects a statistical equilibrium between failure process and repair process in maintained repairable systems that are returned to the operating state following any failure. Roughly, availability may be viewed as the fraction of time that a system is in an operational state independently of how many times it was previously broken and repaired. So, if we may assume that systems components and subsystems are not repairable, we will refer to reliability, while the survivability is a typical feature of restorable systems.

Reliability theory gives the instruments to calculate the reliability parameters of a complex system [1, Dalia M. Elgamel, Prof Dr. Atalla Hashad, 2007], [10, W.D. Gover, 2003], [2, G. Birkin, J. Kennington, E. Olininck, A. Orzynski, and G. Spider, 2005].



The starting point of such evaluation is the availability characterization of the system functional blocks, so, any system functional block can be associated to an instantaneous rate of failure $z(t)$.

Parameters of this model are commonly specified by technology vendors or may be identified by experimental testing before system installation and then may be updated by field measurements during the system operating lifetime.

Any survivability analysis is based on reliability data, i.e. failures rate λ and repair rates μ of the involved equipment. This model is based on the usual approximation of considering a constant failure rate $z(t) = \lambda$, corresponding to a negative exponential reliability function $R(t) = e^{-\lambda t}$. According to such approximation, the mean time to failure $MTTF = 1/\lambda$, and mean time to repair $MTTR = 1/\mu$ respectively, also we introduce the mean time between failures (MTBF) parameter, which is given by $MTBF = MTTF + MTTR$ when failure free operating times and repair times are exponentially distributed [4, Mathieu Clouqueur, Wayne D. Grover, 2002], [6, M. Tornatore, G. Maier, A. Pattavina, 2005]. Since, the links can be called as a repairable links, and the system availability A , its parameter can be called as the average outage.

$$A = \frac{MTBF}{MTBF + MTTR} \quad (1)$$

And it's complementary as the unavailability of a system

$$U = 1 - A \quad (2)$$

To calculate the availability we have to model the network by the Reliability Block Diagram (RBD), as shown in figure (2. b, 3. b)

3. Fiber-Optic Networks Protection strategies:

There are several approaches to ensure fiber network survivability [1, Dalia M. Elgamel, Prof Dr. Atalla Hashad, 2007]. Survivable network architectures are based either on dedicated resources or on dynamic restoration as shown in fig (1). In dedicated-resource protection (which includes automatic protection switching (APS) or self-healing rings), the network resources may be dedicated for each failure scenario, or the network resources may be shared among different failure scenarios. In dynamic restoration, the spare capacity available within the network is utilized for restoring services affected by a failure.

Generally, dynamic restoration schemes are more efficient in utilizing capacity due to the multiplexing of the spare-capacity requirements and provide resilience against different kinds of failures, while dedicated-resource protection schemes have a faster restoration time and provide guarantees on the restoration ability.

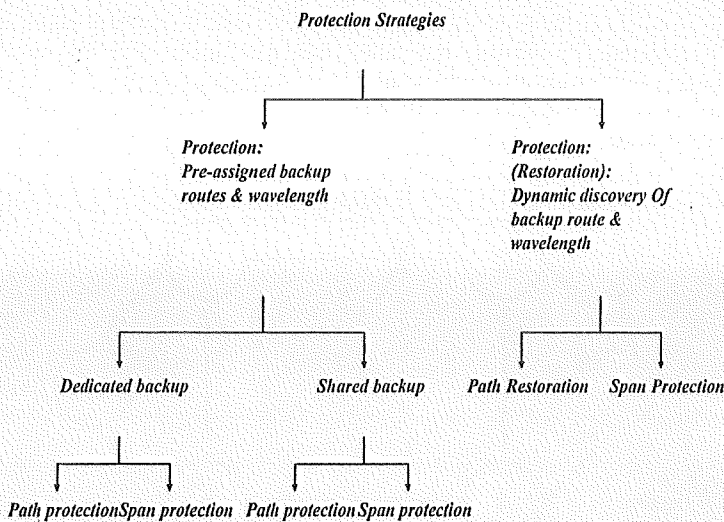


Figure 1: Different schemes for path protection and restoration

Path-protection: is the mechanism that automatically switches the traffic from the working path to a predetermined and diverse path connecting start and end node in case of span or node failure. Path protection can be implemented as 1+1, 1:1, or 1:N. In the case 1+1 traffic is transmitted simultaneously on both path, but one path is selected for transmission usually based on the quality of the signals.

Span-protection: a failed span is bridged by a backup path whose start and end nodes is the adjacent to the failed span. In contrast to path protection it is implicitly assumed no redundant path can be established in the case of node failure.

Shared-Protection: is the mechanism that in shared-path protection, the resources along a backup path may be shared with other backup paths. As a result, backup channels are multiplexed among different failure scenarios (which are not expected to occur simultaneously), and therefore, shared-path protection is more capacity efficient when compared with dedicated-path protection.

Path-restoration: In path restoration, the source and destination nodes of each connection traversing the failed link participate in a distributed algorithm to dynamically discover an end-to-end backup route. If no routes are available for a broken connection, then the connection is dropped, as shown in Fig (2).

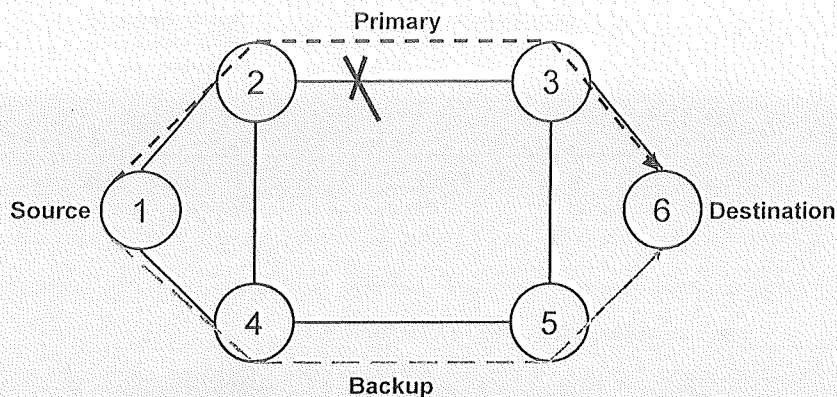


Figure 2a Path protection.

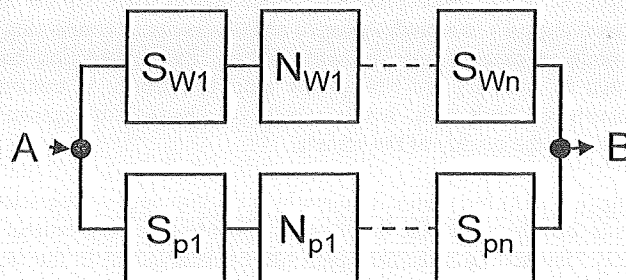


Figure 2b: RBD of path protection.

Span-restoration: In Span-restoration the end nodes of the failed span dynamically discover a route around the link as shown in Fig. (3), all the connections that traverse the failed span are rerouted around that span [3, Hakki C. Cankaya, Ana Lardies, and Gary W. Easter, 2005]. [9, T. H. Shake, B. Hazzard, A. D. Marquis, 2000].

So the basic condition of end-to-end protection is working and protecting light paths being routed along failure independent network. The nodes have been assumed with no failure as if it has a failure according to catastrophic events, also the link-disjoint (the path from the source to destination won't share its protection, fig (3)), it will give absence in failure correlation between any source and destination. So, we can calculate the working path survivability independent from its protection.

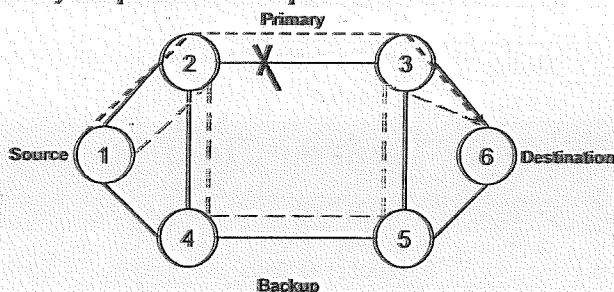


Figure 3 a: Span restoration network.

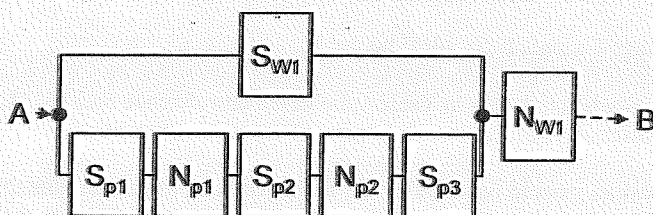
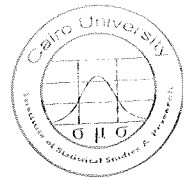


Figure 3 b: RBD of span restoration.

4. Fiber-Optic Networks Design procedure:

The network is modeled by directed graph (digraphs), whose vertices represent the network switching nodes and whose directed edges represent the transmission fibers.



All networks considered have a pair of unidirectional working fibers (constituting a bidirectional working span) and a pair of unidirectional protection fibers (constituting a bidirectional protection span).

The C language is used for simulating the model, and the minimum-weight path-searching algorithm, or Dijkstra algorithm to find the highest unavailability along the whole network, so the search can define the weakest points on it to show how to give the maximum protection for it. Also, the program has Breadth-First Search algorithm to find alternate paths when a single cut happens, according to the remaining capacity.

Table 1: Failure rates and repair time for optical components.

Module	Failure Rate (R)	MTTR
Fiber	2.12566E-07	(12),(12) hours
Regenerator	3.35521E-06	(2),(2) hours
Tx, Rx	3.35521E-06	2
Amp	(4.22508E-06), (2000FIT*)	2,
Node	10.0E-6	6
Average Equipment Failure		2 hours

* FIT: Average Failure rate in $10E^09$ hours.

Analytical calculations given in Table 1 are based on reference data [5, M. Held, L. Wosinska, p. M. Nellen, C. Mauz, 19877- 9] for illustration.

Where the values used in equation (2) as, failure rates λ_N and mean time to repair $MTTR_N$ for nodes, failure rate per kilometer λ_{Fiber} of fiber in the optical cable, mean time to repair $MTTR_S$ of spans, and the failure rate of optical amplifiers λ_{OA} . $MTTR_S$ applies to the complete span, i.e. repair times of optical amplifiers and fiber optical cables are considered equal. We have assumed that bidirectional Optical Amplifiers with a failure rate λ_{OA} are located every 50 kilometers on a span which yields an optical amplifier spacing constant $C_{OA} = 0.02$ per kilometer. The failure rate of a span λ_S is therefore its length L multiplied by the sum of fiber failure rate per kilometer λ_{Fiber} and optical amplifier failure rate per kilometer. Figure 4 shows the difference between path and span:

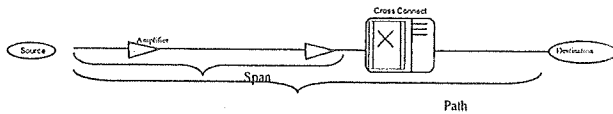


Figure 4: End-to-end path.

$$\lambda_s = L \cdot (\lambda_{\text{fiber}} + C_{OA} \cdot \lambda_{OA}) \quad (3)$$

A node containing the necessary line and cross connecting equipment is represented by a single block characterized by λ_N and $MTTR_N$. In order to compare the different protection strategies calculations are performed for a span length $L_{\text{long}} = 904\text{km}$ representing an average value in a long haul network, i.e. a pan European network, and $L_{\text{short}} = 91\text{km}$, i.e. an average value in metropolitan network for comparison.

In case of no protection the RBD of a particular connection is a simple series structure consisting of n spans and $n-1$ nodes with the availability expressed as: [5, M. Held, L. Wosinska, p. M. Nellen, C. Mauz, 1987].

$$A = \left(\frac{\mu_S}{\lambda_S + \mu_S} \right)^n \cdot \left(\frac{\mu_N}{\lambda_N + \mu_N} \right)^{n-1} \quad (4)$$

Survivability for generalized path protection

$$A = 2 \cdot \left(\frac{\mu_S}{\lambda_S + \mu_S} \right)^n \cdot \left(\frac{\mu_N}{\lambda_N + \mu_N} \right)^{n-1} - \left(\frac{\mu_S}{\lambda_S + \mu_S} \right)^{2n} \cdot \left(\frac{\mu_N}{\lambda_N + \mu_N} \right)^{2n-2} \quad (5)$$

For generalized span protection

$$A = \left(1 - \frac{3\lambda_S^2}{\mu_S^2} - \frac{2\lambda_S\lambda_N}{\mu_S\mu_N} \right)^n \cdot \left(\frac{\mu_N}{\lambda_N + \mu_N} \right)^{n-1} \quad (6)$$

We use these three equations to calculate survivability in different conditions [1, Dalia M. Elgamel, Prof Dr. Atalla Hashad, 2007].

5. Fiber-Optic Networks Modeling Results:

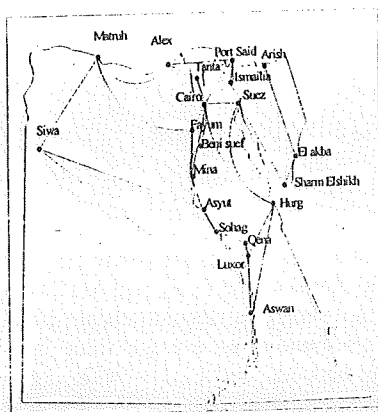


Figure 5: Egypt optical transport network

A proposed network between the 20 biggest Egyptian cities (nodes), are shown in Fig. 5. The distance range 90 Km to 907 Km, the distance between amplifiers is kept constant (50 Km). The data in table 1 are substituted in equations 4, 5, and 6. The unavailability of the networks is shown in Figures 6, 7, and 8 for different distances. When using 90 Km in the equations, span protection is better than path protection as shown in Figure 6.

This is due to span failure rates are low, also spans have redundancy where as nodes of working path represent single point of failures in the case of span protection, this is not affected for longer spans.

For larger distances and number of nodes the path protection is preferred than span protection because of the multiple routes available through the whole network. However span protection is performing better when the network contains a smaller number of nodes as shown in Figure 7.

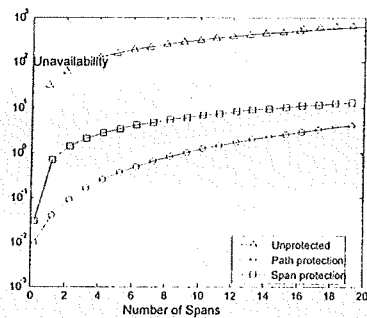


Figure 6: Unavailability for short distances 90 Km

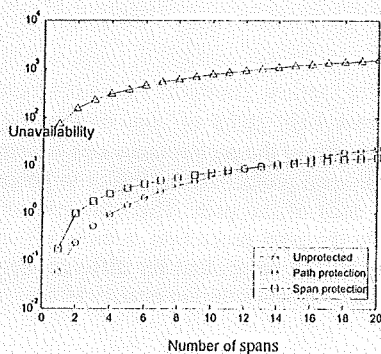


Figure 7: For moderate distances 224 Km.

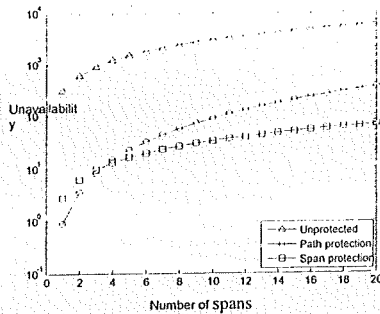
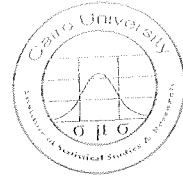


Figure 8: Unavailability for long distances 904 Km.

We can see for long distances path protection is preferred than span protection.

6. Conclusions:

The survivability models for Optical Transport Networks have been presented, and a purposed network light paths of 20th biggest Egyptian cities is studied. The simulation results of the performance can be improved by applying the path protection or span protection according to geographic distribution of the nodes at the ends of the path. The span protection is preferred for short distance while the path protection is more efficient in long distances. The algorithms used can be applied for different networks design, for both availability calculations and solution qualification.



References

- [1] Dalia M. Elgamel, Prof Dr. Atalla Hashad, 2007, "Survivability Design and Protection Strategies for a Proposed Egyptian Backbone" MSC Thesis Computer Engineering Department, College of Engineering & Technology, Arab Academy for Science & Technology & Maritime Transport Feb 2007
- [2] G. Birkin, J. Kennington, E. Olininck, A. Ortynski, and G. Spider, 2005, "Design Strategies for Meeting Unavailability Targets Using Dedicated Protection DWDM Networks " Available at <http://engre.smu.edu/mjlk/publications.html>.
- [3] Hakki C. Cankaya, Ana Lardies, and Gary W. Easter, 2005, "Survivability- aware Analysis and Evaluation of Mesh and Ring Architectures for long-haul Networks", Available at <http://www.Scs.org/getDoc/d=1643>.
- [4] Mathieu Clouqueur, Wayne D. Grover, 2002. "Survivability analysis of Span-Restorable Mesh Networks", IEEE J. Sel. Areas Comm. Vol.20, NO.4, pp.810-821, May 2002.
- [5] M. Held, L. Wosinska, p. M. Nellen, C. Mauz, 1987, "Consideration of connection survivability optimization in optical networks", this work was partly financed by the Swiss Federal Office for Education and Science BBW (project C01.1987).
- [6] M. Tornatore, G. Maier, A. Pattavina, 2005. "Survivability Design of Optical Transport Network", IEEE J. Selected Areas Comm., Vol.23, NO.8, pp.1520-1532, August 2005.
- [7] M. To, and P. Neusy, 1994, "Unavailability of Long Haul Networks", IEEE J. Selected Areas Comm., Vol.12, NO.1, pp100-110, January 1994.
- [8] S. Ramamurthy, L. Sahasrabuddhe, Biswanath Mukherjee, 2003, "Survivable WDM Mesh Networks", IEEE J. Light Wave Tech., Vol. 21, No.4, April 2003.
- [9] T. H. Shake, B. Hazzard, Á, D. Marquis, 2000. "Assessing Network Infrastructure Vulnerabilities to Physical Layer Attacks", this work was sponsored by the Defense Advance Research Projects Agency under Air Force Contract F19628-95-C-0002
- [10] W.D. Gover, 2003. "Mesh-Based Survivable Networks", Publisher: Prentice Hall PTR, August 26, 2003.



Cairo University

Institute of Statistical Studies and Research

**The 42nd Annual Conference on Statistics,
Computer Sciences and Operation Research
2-5 Dec, 2007**



An Optimization Energy Model for Egypt

A.A. Tharwat*, M.M. Saleh*, D.S. Ali*

* Decision Support Department, Faculty of Computers and Information, Cairo University

Abstract: Energy plays an essential role in peace and war for any country. The Egyptian energy problem which represented in the following question: how could the limited available resources from the conventional energy (Petroleum) cover the high growth of demand of energy? The Egyptian energy problem erupted during the war in October 1973; the continuity of the problem gave a good motivation to the following two approaches: (1) creating optimal models to redistribute the available energy resources and (2) to search for alternative non-conventional energy resources (Natural Gas...etc). During the last three decades there are many energy models were developed to solve the energy problems [1], [3], [4], [5]. This article introduces a new mathematical energy model (Distribution of Energy Sources on Different Sectors DESDS), three objectives are considered in the proposed multi-objective nonlinear mathematical model: the 1st is to minimize the total cost, the 2nd is to minimize the pollution ratio, and the 3rd is to minimize the total governmental subsidy. Finally the model is implemented by using a real life data obtained from the Egyptian Organization of Energy Planning (EOEP), the results were compared with the EOEP distribution plan to explain the benefits of applying the proposed model to those types of problems.

Keywords: Energy Models, Mathematical Models, and Multiobjective Nonlinear Programming.

تلعب الطاقة دورا أساسيا في حياة أي دولة سواء أثناء الحرب أو السلام. بدأت مشكلة الطاقة في مصر، والتي تتمثل في التزايد المستمر للكميات المطلوبة من الطاقة التقليدية (البترول) مع محدودية المصادر التي تغطي تلك الطلبات، تجذرت مشكلة الطاقة التقليدية عام ١٩٧٣ في مصر، واستمرارية تلك المشكلة أعطت دوافع قوية للتقدم في اتجاهين بحثيين على التوازي، أولهما استنباط النماذج الرياضية التي تبحث للتوزيع الأمثل لمصادر الطاقة المحدودة، والثاني يتمثل في المحاولة للبحث عن مصادر بديلة للطاقة غير التقليدية (الغاز الطبيعي،...). في العقود الثلاثة الماضية تم تطوير العديد من النماذج الرياضية التي اختتمت بحل مشكلة الطاقة، ويقدم هذا البحث نموذج رياضي جديد (DESDS) وتم التعامل مع ثلاثة أهداف في النموذج متعدد الأهداف غير الخطية وهي: (١) تخفيض التكلفة الكلية لتوزيع الطاقة، (٢) تخفيض نسبة التلوث، (٣) تخفيض الدعم الحكومي للمصادر المخصصة لتوزيع الطاقة. كما تم تطبيق النموذج المقترح على بيانات حقيقية حصلنا عليها من جهاز تخطيط الطاقة بمصر، وقورنت النتائج التي استخرجت من النموذج مع ماسبق استخدامه من جهاز تخطيط الطاقة لتوضيح مزايا استخدام النموذج في تلك النوعية من المشاكل.



Introduction.

Sufficient supplies of energy have always played a crucial role for people on the earth. In the early days heat was the most important energy form, providing people with cooking facilities and keeping them warm. Today most human activities are concerned with use of various forms of energy, and we experience a growth in energy consumption in most parts of the world. At the same time, the knowledge and concern about environmental consequences of energy utilization are increasing. So it is needed to find better energy supply systems for the future. Different computer models were created to improve energy supply systems. Also there are several types of modeling tools were used such as optimization techniques, simulation approach, and the system dynamic concepts.

Optimization techniques are important in many fields of science and engineering. A lot of modeling, design, controls and decision-making problems can be formulated in terms of mathematical optimization. Moreover, there are several mathematical optimization techniques such as linear programming (LP), nonlinear programming (NLP), dynamic programming (DP), and multiobjective (linear / nonlinear) programming (MOP).

Many real-world decision making problems need to achieve several objectives such as: minimizing risks, maximizing reliability, minimizing deviations from desired levels, minimizing cost, etc. The main goal of single- objective (SO) optimization problem is to find an optimal solution, which corresponds to the minimum or maximum value of the single objective function that lumps all different objectives into one. This type of optimization problems is useful as a tool, which should provide decision makers with insights into the nature of the problem, but usually cannot provide a set of alternative solutions that trade different objectives against each other. On the other hand, there is no single optimal solution in a multiobjectives optimization problem with conflicting objectives. The interaction among different objectives gives rise to a set of compromise solutions, largely known as the trade-off, non-dominated, non-inferior or Pareto-optimal solutions [12]. There are several methods, which can handle and deal with multiobjectives problems such as: *weighting objectives method, hierarchical optimization method, trade-off method, global criterion method, and distance functions method.*

1. Modeling Approaches:

Models differentiate according to the model designer view and the model field (e.g. economic, engineering...etc). Energy system models are often classified mainly into the following three types [4]:

- A. System engineering Models,
- B. Macroeconomic Models, and
- C. System Dynamics Models.



System engineering models interest in technical engineering side of energy problem and discuss in details the different energy conversion technologies. Moreover, this type of models concentrate on the demand side, which calculate the energy demand in a separate model, and treat the different sorts of energy demand as a fixed parameter within the energy system model. For example, MARKAL model (optimization model---linear/dynamic programming) was developed by the Energy Technology Systems Analysis Program (ETSAP) of the International Energy Agency (IEA) [5], and EFOM-ENV model (optimization model---linear/dynamic programming) was developed by European Commission DDG-XII F/1 (Belgium) [3].

Macroeconomic models tend to focus on the demand side. These models usually incorporate larger parts of the economy. The development of each sub-sector in the economy depends on sensitivities, or Elasticities, between relevant economic variables within and between the sub-sectors. The interaction between different sub-sectors can be treated using so-called input-output analysis, which is a model description of how commodities flow between sectors. Elasticities and input-output coefficients are calculated using regression analysis based on historical data. The energy supply side in the system is given less attention and is often limited to include a description of average investment and operating costs of different supply technologies.

An underlying assumption in macroeconomic models is that *the state of the economy is in equilibrium*. It can be argued that this is a questionable assumption, because the rare occurrence of this case in the real life situations. At least it takes some time to reach equilibrium from the starting point. The high dependence of historical data in macroeconomic is also sometimes subject to critics, because historical relationships might not always be valid in the future. For example, ENPEP model was developed by International Atomic Energy Agency (IAEA), Austria, and ENERPLAN model (economic and simulation model) was developed by Tokyo Energy Analysis Group (Japan) [3].

A third type of models which used in energy system models is the of system dynamics. It is a tool for developing general mathematical simulation models for different kinds of processes (e.g. technical, biological or economic processes). Stocks, flows and feedback loops are the building blocks in a system dynamics model. System dynamics models aren't dependent on historical data as macroeconomic models, and can also simulate those systems that are not in equilibrium.

It is important to differentiate between optimization and simulation models. An optimization model tries to find the best possible solution to a specified problem based on maximization or minimization of an objective function. Optimal production or system expansion plans are often found using optimization models. Linear, Nonlinear, Dynamic programming are optimization techniques that are very often used in system engineering and Macroeconomic models. Simulation models try to predict the behavior of a system based on a set of assumptions, without necessarily searching for optimal solutions. System dynamics models are based on simulation models.

Traditionally total energy system optimization models have often been applied in energy planning on the national level. However, the international trend of deregulating energy markets reduces the relevance of direct national energy planning.

2. Problem Definition:

Energy is a driving engine for development. It is one of the basic factors to improve quality of life and to achieve the goals of the sustainable development. The policies of the electricity and energy sector aim at the optimum utilization of the primary energy resources. This will ensure timely supply of electrical energy in capacities sufficient of foster the economic and social development. The demand for energy has increased tremendously to meet technological needs and changing lifestyles as well as population growth. Conventional energy resources are reaching their natural limits and development of new and renewable sources of energy has not yet reached an economically feasible level permitting their use as an alternative to conventional resources. Figure-1 will explain simplify give the energy problem.

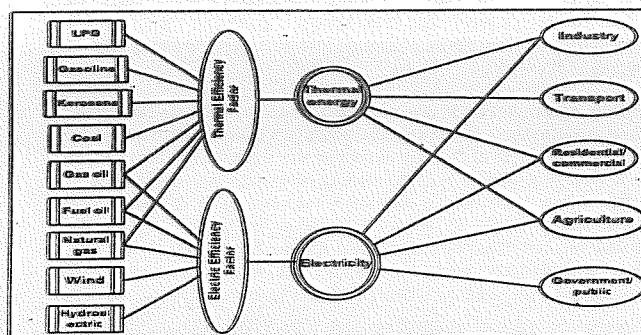


Figure -1 defines the energy problem

Energy Sources:

- Gasoline, Kerosene, Coal oil, and Butagas (LPG) give thermal energy to different sectors.
- Gas oil, Fuel oil, and Natural gas give both electricity and thermal energy to different sectors.
- Wind and Hydropower gives electricity energy to the electric-sector only.

Energy demand of different Sectors:

- Industrial-sector, Agriculture-sector, Transportation-sector, and Residential & commercial sector: need both electricity and thermal energy. (Electricity could be neglected according to annual report of Electricity Corporation for the Transportation-sector)
- Government & public sector and Electric-sector: need only electricity energy.

Energy source can't be used in any purpose when it is in a crude form. Each source must pass through certain intermediate process. Some of source is lost during this process as water vapor, deposits, or other types. For example, in case of natural gas, the production amount of thermal energy is only 90% from the consumption amount of natural gas. This percentage is called efficiency factor (see table-2) [8]. There are efficiency factor for thermal energy and electricity.

Source	Efficiency factor to convert source to electric energy %	Efficiency factor to convert source to thermal energy %
Petroleum product	26.6	90
Coal	0	70
Natural gas	17.5	90
Wind	30.7	0
LPG	0	90
Hydroelectric	95	0

Table 2: shows efficiency factor for conversion of energy sources to final energy form

Pollution Problem: recently the protection and preservation of the environment is one of the most important issues for all scientists and specialists interested in analyzing the interactions between energy and environment and their impact on sustainable development (see figure-2).

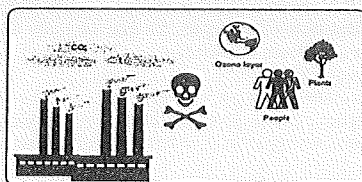
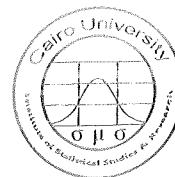


Figure 2: shows dangerous effects of CO₂

Naturally, the concern is to minimize the adverse impacts of production and use of energy on the environment. Increasing concentrations of carbon dioxide (CO₂) emissions are the product of burning fossil fuels such as petroleum products, natural gas, and coal. Exploration, production, manufacturing, transportation and processing these fuels all contribute to (CO₂) emissions. Table-3 shows emission factors of CO₂ by fuel oil [7] [9].

Fuel Type	Ton Carbon Dioxide/ petroleum ton equivalent
LPG	2.9837
Gasoline	3.1046
Kerosene	3.2160
Gas oil	3.2093
Fuel oil	3.2094
Coal	2.9473
Natural gas	1.6115

Table 3: Emission Factors of CO₂ by Fuel Oil



From above, an energy problem is considered from the big problems, which can be handled with a lot of methods according to the field of handler. Furthermore, this article will concentrate on a distribution process of energy sources on different sectors. DESDS model searches about optimum distribution of energy sources on different sectors, in addition to minimize the total cost, the pollution ratio (emissions of CO₂ by fuel oil), and governmental subsidy on different energy sources under some manufacturing constraints. Next, DESDS model will be interpreted in details.

3. Proposed Model (DESDS model):

Energy system planning requires the use of planning tools. The mathematical models of real-world energy systems are usually linear optimization programs. But, DESDS model is designed by using the multiobjective nonlinear programming. There are some remarks in DESDS model for simplifying understanding the model idea:

1. Thermal energy sources, such as LPG, gasoline, kerosene, coal, gas oil, fuel oil and natural gas can provide thermal energy. Electrical energy sources such as gas oil, fuel oil, natural gas, wind and hydroelectric can provide electricity. Moreover, some energy sources can provide electricity and thermal energy such as gas oil, fuel oil and natural gas.
2. The gathered data aren't in suitable form because of the variations of units of different sources, such as: petroleum products are measured with metrics tons and natural gas is measured with cubic meters. So that, we unify all the units in one general unit which is petroleum ton equivalent (PTE). Also Cost is evaluated in L.E/ petroleum ton equivalent.
3. Electrical sector is considered as supply element and demand element at the same time, because it supplies other sectors with electricity according to their needs, and it demands electrical energy sources to be converted to electricity.

3.1 Assumptions of DESDS Model:

There are some assumptions which are needed for building the model, as follows:

1. Thermal energy sources have equal opportunity in satisfying the needs of different sectors.
2. Electric sector is considered a unique source of electricity.

3.2 Index of DESDS Model:

Sets of model are divided mainly into two basic sets. The set I represents an index for energy sources $\{i = 1, 2, \dots, 9\}$, and it is divided into three subsets I_1 , I_2 , and I_3 :

- I_1 : Represents energy sources which can give electricity.
- I_2 : Represents energy sources which can give thermal energy.
- I_3 : Represents energy sources which can provide both thermal and electrical energy.

And the set J stands for an index for different sectors ($j = 1, 2, \dots, 5$), which also divided into two subsets J_1 and J_2 .

J_1 : Represents sectors which need thermal energy.

J_2 : Represents sectors which need electricity.

3.3 Parameters and Decision Variables of the DESDS:

3.3.1 Model Parameters:

There are several parameters in DESDS model. Most of data are gotten from the annual report of EOEP and a few of data is estimated.

DT_j	$(j \in J_1)$	Represents the domestic demand of the j^{th} sector for thermal energy sources.
DE_j	$(j \in J_2)$	Represents the domestic demand of the j^{th} sector for thermal electricity.
TC_i	$(i \in I_1)$	Represents purchasing cost of the i^{th} thermal energy source.
EC_j	$(j \in J_2)$	Represents purchasing cost of the j^{th} sector for electricity.
SF_i	$(i \in I)$	Represents subsidy factor of the i^{th} energy source.
RE_i	$(i \in I_1)$	Represents electrical efficiency factor of the i^{th} energy source.
RT_i	$(i \in I_2)$	Represents thermal efficiency factor of the i^{th} energy source.
RTE_i	$(i \in I_3)$	Represents electrical efficiency factor of the i^{th} energy source, which can provide both thermal energy and electricity.
FP_i	$(i \in I)$	Represents emission factors of CO_2 of the i^{th} energy source.
MP_i	$(i \in I)$	Represents maximum production of the i^{th} energy source.

3.3.2 Model Variables:

Some of model decision variables represent proportions and have percentage values between $[0-1]$ such as TS_{ij} , ES_i , and ELS_i , while the other of variables are positive numbers.

DS_i	$(i \in I)$	Represents domestic supply of different energy sources.
TS_{ij}	$(i \in I_1) (j \in J_1)$	Represents sector share of total production of thermal energy sources
ES_i	$(j \in J_2)$	Represents sector share of total production of electricity.
TEP	-----	Represents total production of electricity after the conversion process. It is found in two constraints, where it is updated with its value by the first constraint (represent the actual amount of electricity) to know the permitted available amount of electricity production and the second constrain distributes this amount of electricity on several sectors to satisfy its needs from electricity.

EIS _i	(i ∈ I ₁)	Represents a share of electric sector from energy sources, which can provide both thermal energy and electricity. This variable can be classified as auxiliary variable.
------------------	-----------------------	--

3.4 Formulation of DESDS Model:

In this section the constraints and objective functions of DESDS model will be discussed and interpreted in details according to our previous knowledge with the energy problem, index of model, parameters, and decision variables.

1. Model Constraints:

In general, model constraints mean that a restriction on objective functions values. Furthermore, objective functions can't achieve without satisfying all constraints. Next, model constraints will be discussed in details, as follows:

1. To satisfy the sectors' demand from thermal energy sources, the domestic demand of the j^{th} sector for thermal energy sources - DT_j (for each $j \in J_1$) should be less than or equal the sum of shares from total production of thermal energy sources (i.e. the shares for each $i \in I_1$). Figure-3 illustrates the satisfaction concept. The mathematical form is given as follows:

$$\sum_{i \in I_1} DS_i * RT_i * TS_{i,j} \geq DT_j \quad (\text{for each } j \in J_1)$$

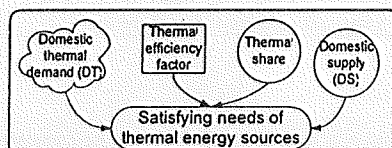


Figure-3 shows an influence diagram of distributing thermal energy (the square form represents input parameters, cloud form represents exogenous variables (parameters), circle form represents endogenous variables (decision variables), and oval form represents the overall performance [10]).

2. To satisfy the sectors' demand from electricity, the domestic demand of the j^{th} sector for electricity - DE_j (for each $j \in J_2$) should be less than or equal the sum of shares from total production of electricity. Figure-4 illustrates the satisfaction concept. The mathematical form is given as follows:

$$TEP * ES_j \geq DE_j \quad (\text{for each } j \in J_2)$$

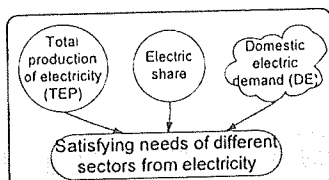


Figure 4: shows an influence diagram of distributing electricity

3. Some energy sources can provide both thermal energy and electricity. Furthermore, this constraint aims to guarantee that the sum of shares from total consumption amount of i^{th} energy sources (in both the two purposes) are less than or equal the available production amount of it (i.e. the shares for each $i \in I_3$). From above, ELS_i , $TS_{i,j}$ have values between 0 and 1. The mathematical form is given as follows:

$$ELS_i + \sum_{j \in J_1} TS_{i,j} \leq 1 \quad (\text{for each } i \in I_3)$$

4. To satisfy the electric sector' demand from electrical energy sources, the total production of electricity by electric sector – TEP should be less than or equal the sum of shares from total production of electrical energy sources (i.e. the shares for each $i \in I_1$). Figure-5 illustrates the satisfaction concept. The mathematical form is given as follows:

$$\sum_{i \in I_1} (ELS_i * DS_i * RTE_i) + (RE_{(wind)} * DS_{(wind)} + RE_{(hydroelectric)} * DS_{(hydroelectric)}) - TEP \geq 0$$

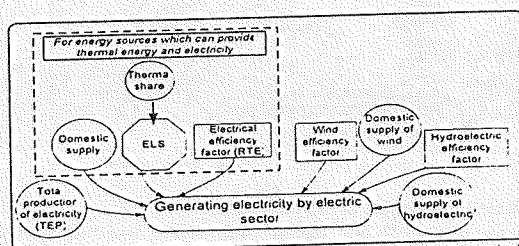


Figure 5: shows an influence diagram of generating electricity by electric sector (the hexagon form represents auxiliary variables)

5. To guarantee that the consumption amount of i^{th} thermal energy source (for each $i \in I_3$) doesn't violate the total production of it, the sum of shares from total production of thermal energy sources (i.e. the shares for each $j \in J_1$) is less than or equal 1. The mathematical form is given as follows:

$$\sum_{j \in J_1} TS_{i,j} \leq 1 \quad (\text{for each } i \in I_3)$$

6. To also guarantee that the consumption amount of electricity doesn't violate the total production of it, the sum of shares from total production of electricity (i.e. the shares for each $j \in J_2$) is less than or equal 1. The mathematical form is given as follows:

$$\sum_{j \in J_2} ES_j \leq 1$$

7. To restrict the consumption amount of different energy sources, the domestic supply of i^{th} energy source $-DS_i$ (for each $i \in I$) is less than or equal the maximum production $-MP_i$ of it. The mathematical form is given as follows:

$$DS_i \leq MP_i \quad (\text{for each } i \in I)$$

8. To void negative values for the values of decision variables, all variables should be greater than or equal zero. In addition, this constraint is called negativity constraint.

$$DS_i, TEP, TS_{i,j}, ES_j, ELS_i \geq 0$$

2. Objective Functions:

From above, DESDS model has mainly three objective functions as follows:

1. Objective function of total cost:

Each sector needs thermal energy sources or/and electricity. Furthermore, this objective function aims to minimize purchasing cost of energy sources on different sectors. Figure -6 illustrates the objective function of total cost.

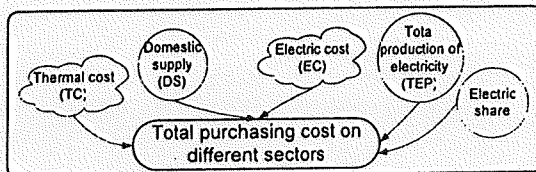


Figure 6: shows an influence diagram of total cost objective function

Objective function of total cost consists of two basic parts:

- firstly, purchasing cost of PTE of i^{th} thermal energy source for satisfying needs of different sectors:

$$\sum_{i \in I_1} TC_i * DS_i$$

- Secondly, purchasing cost of PTE of electricity for satisfying j^{th} sector' demand:

$$\sum_{j \in J_2} EC_j * TEP * ES_j$$

The final mathematical form is given as follows:

$$F_1: \sum_{i \in I_2} TC_i * DS_i + \sum_{j \in J_2} EC_j * TEP * ES_j$$

2. **Pollution objective function:** aims to minimize the dangerous of CO₂ emissions of i^{th} energy source (for each $i \in I$) for protecting the environment. Figure-7 illustrates the main idea of this objective function. The mathematical form is given as follows:

$$F_2: \sum_{i \in I} FP_i * DS_i$$

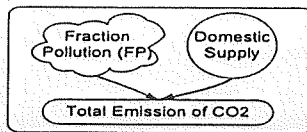


Figure7: shows an influence diagram of pollution objective function

3. Subsidy objective function:

The government is presenting a lot of subsidy on different energy sources for reducing its price on a domestic consumer. As a matter of fact, subsidy money is considered as load on the government and the national income. For these reasons, subsidy objective function aims to minimize the subsidy government for i^{th} energy source (for each $i \in I$). The mathematical form is given as follows:

$$F_3: \sum_{i \in I} SF_i * DS_i$$

3.5 General Formulation of DESDS Model:

Objectives:

$$\text{Minimize } \sum_{i \in I_2} TC_i * DS_i + \sum_{j \in J_2} EC_j * TEP * ES_j$$

$$\text{Minimize } \sum_{i \in I} FP_i * DS_i$$

$$\text{Minimize } \sum_{i \in I} SF_i * DS_i$$

Subject to

$$\sum_{i \in I_2} DS_i * RT_i * TS_{i,j} \geq DT_j \quad (\text{for each } j \in J_1)$$

$$TEP * ES_j \geq DE_j \quad (\text{for each } j \in J_2)$$

$$ELS_i + \sum_{j \in J_1} TS_{i,j} \leq 1 \quad (\text{for each } i \in I_3)$$

$$\sum_{i \in I_1} (ELS_i * DS_i * RTE_i) + (RE_{(wind)} * DS_{(wind)} + RE_{(hydroelectric)} * DS_{(hydroelectric)}) - TEP \geq 0$$

$$\sum_{j \in J_1} TS_{i,j} \leq 1 \quad (\text{for each } i \in I_2)$$

$$\sum_{j \in J_2} ES_j \leq 1$$

$$DS_i \leq MP_i \quad (\text{for each } i \in I)$$

$$DS_i, TEP, TS_{i,j}, ES_j, ELS_i \geq 0$$

3.6 Input Data: [6]

Now after explaining DESDS model in details, we can use a real life data to implement the model. The main source of input data is Egyptian Organization of Energy Planning (EOEP). In addition, this input data is the values of DESDS model parameters. The input data (the parameters of DESDS model) will be given, as follows:

1. In DESDS model, the cost is divided into two main types:

- Thermal cost (TC): represents purchasing cost of PTE of thermal energy sources (see, table-4).

Source	Thermal cost (000*L.E/ 000'petroleum ton equivalent)
LPG	225
Gasoline	246.9
Kerosene	234.94
Coal	225.987
Gas oil	226.4
Fuel oil	176.904
Natural gas	150.651
Wind	----
Hydroelectric	----

Table 4: represents the cost of thermal energy sources

- Electrical cost (EC): represents a purchasing cost of electricity from electric sector. Moreover, the sectors get the electricity with different costs. In real life, electricity is evaluated by kilowatt hours (k.w.h), where the one petroleum ton equivalent = 4474.273 k.w.h (see, table-5).

Sector	Electric cost (000'L.E / 000'petroleum ton equivalent)
Industry	593.74
Transport	----
Residential/Commercial	834.004
Agriculture	512.75
Government/Public	100.895

Table 5: represents electric cost

2. Available and maximum amount (MP_i) of different energy sources will be given in (table-7).

Source	Available amount of energy sources	Maximum amount of energy sources
LPG	1,864,625	2,613,625
Gasoline	1,360,580	2,500,580
Kerosene	1,109,800	1,536,800
Coal	39,000	45,000
Gas oil	5,892,848	8,000,848
Fuel oil	5,261,404	661,404
Natural gas	18,390,000	19,390,000
Wind	12,000	12,000
Hydroelectric	2,997,000	5,997,000

Table 6: represents the available and maximum amount of different energy source

3. Some sectors need thermal energy, electricity or both. The sectors' demands of thermal energy sources and electricity will be given in table-7.

Sector	Thermal Demand (DT _i)	Electricity Demand (DE _i)
Industry	11622658	2088000
Transport	9634690	0
Residential/Commercial	4071075	2162000
Agriculture	845000	212000
Government/Public	0	18210074

Table 7: represents the sectors' demands of thermal energy sources and electricity

4. The government subsidy factor on different energy sources will be given in table-8.

Source	Subsidy Factor (SF _i)
LPG	80
Gasoline	70.2
Kerosene	75.1
Coal	100.91
Gas oil	76.12
Fuel oil	78.2
Natural gas	100.3
Wind	100
Hydroelectric	105.5

Table 8: represents the amount of subsidy

3.7 Output analysis:

The solution method which was used in solving the model will be given in details in the appendix. The optimal values of decision variables are as follows:

1. The values of DS_i variable.

Source	Domestic supply
LPG	1,864,625
Gasoline	2,500,580
Kerosene	1,536,800
Coal	39,000
Gas oil	8,000,848
Fuel oil	6,661,404
Natural gas	19048.019
Wind	12,000
Hydroelectric	5,997,000

Table 9: represents the values of DS_i variable

2. The share of each sector from thermal energy sources will be shown in table-10. These values were resulted by multiplying the DS_i and TS_{ij} for each source and each sector.

Source	Industry	Transport	Agriculture	Residential & commercial	Electric
LPG	1,754,612.125	-----	-----	110,012.875	-----
Gasoline	-----	2,500,580	-----	-----	-----
Kerosene	-----	974331.2	562,468.8	-----	-----
Coal	39,000	-----	-----	-----	-----
Gas oil	-----	1,240,131.44	376039.856	-----	5896,624.976
Fuel oil	-----	5995263.6	-----	-----	699447.42
Natural gas	11,124,043.096	-----	-----	4,419,140.408	4190564.18

Table 10: represents the share of each sector from thermal energy sources

3. Total production of electricity is the value of TEP variable, where TEP = 8,191,067 petroleum ton equivalent. Share of each sector from electricity are be given in table-11. These values were resulted by multiplying the DS_i and ES_i for each source and sector.

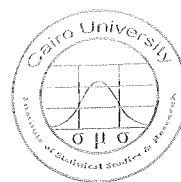
Sector	Share of electricity
Industry	2,088,722.1
Residential & commercial	2,162,442.8
Agriculture	2,121,486.4
Government & Public	1,818,416.9

Table 11: represents the share of each sector from electricity

4. Consumption amount of electrical energy sources for producing the electricity in table-12.

Source	Electric sector
Gas oil	5,760,610.56
Fuel oil	795,198.308
Natural gas	4,265.800
Wind	12,000
hydroelectric	5,997,000

Table 12: represents the consumption amount of electrical energy sources



5. The values of the three objective functions are:

$f_1 = 11,580,570,000$ L.E,

$f_2 = 96,136,190$ emissions of CO_2 ,

$f_3 = 4,145,052,000$ L.E,

Where the

TEP = 8,191,067 petroleum ton equivalent.

Conclusion:

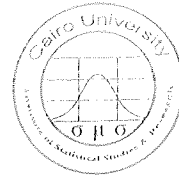
This article aims mainly to design mathematical energy model (DESDS). DESDS model could be formulated by using multiobjective nonlinear programming. The model was implemented by using a real life data from EOEP. After testing DESDS model under different scenarios (in, appendix), DESDS model achieved results better than the annual report of EOEP in all cases, as follows:

1. Reducing the usage of petroleum products.
2. Decreasing the total cost by %10.4706.
3. Achieving more protection for the environment by decreasing emissions of CO_2 by %11.6512.
4. Satisfying sectors' needs from different energy sources.
5. Decreasing the governmental subsidy by % 1.8912.

For the purpose of certain from goodness of the model, the model was tested in its worst cast (in, appendix). DESDS model in its worst case achieved also results better than EOEP.

References

- [1] Ad J. Seebregts¹, Gary A. Goldstein², Koen Smekens¹, "Energy/Environmental Modeling with the MARKAL Family of Models", 1 Energy Research Centre of the Netherlands (ECN), Policy Studies unit, P.O. Box 1, 1755 ZG Petten, The Netherlands, e-mail: seebregts@ecn.nl 2 International Resources Group Ltd., 1211 Connecticut Ave. NW, Suite 700, Washington, DC 20036, USA, e-mail: ggoldstein@irgltd.com
- [2] Anthony Brooke, David Kendrick, Alexander Meeraus, and Ramesh Raman, "GAMS A User Guide", Development Corporation, December 98GAMS.
- [3] Beeck. N. V., "Classification of Energy Models", Communicated by dr.ing. W. van Groenendaal, Tilburg University & Eindhoven University of Technology, May 1999. <http://greywww.w.kub.nl:2080/greyfiles/fcw/1999/doc/777.pdf#search=energy%20models%20777>
- [4] Botterud. A., "Energy System Models", NTNU Norwegian University of Science and Technology, Faculty of Electrical Engineering and Telecommunications, Department of Electrical Power Engineering, January 2001.



- [5] David L. Greene, "Long-Term Energy Scenario Models: A Review of the Literature and Recommendations", National Transportation Research Center Oak Ridge National Laboratory, February 28, 2001.
[http://www.ott.doe.gov/pdfs/models2050.pdf#search=models%202050energy'](http://www.ott.doe.gov/pdfs/models2050.pdf#search=models%202050energy)
- [6] EOEP (Egyptian Organization of Energy Planning). Annual report, (2000/2001).
- [7] ETSAP (2001) ETSAP website: http://www.ecn.nl/unit_bs/etsap
- [8] Hassan. Y, "Optimum Distribution of Energy Sources on Different Sectors". Military Technical College (1980).
- [9] IPCC (Intergovernmental Panel on Climate Change), GHG Inventory Guidelines, 1996.
- [10] John Wiley and Sons, "System and Decision Making" in a management science approach, *university of Canterbury, Christchurch, New Zealand, May 1995*.
- [11] Ralph E. Steuer, "Multiple Criteria Optimization: Theory, Computation, and Application", College of Business Administration, University of Georgia.
- [12] Tobias Canz, "Fuzzy Linear Programming in DSS for Energy System Planning", IIASA "International Institute for Applied Systems Analysis" A-2361 Laxenburg Austria, E-Mail: info@iiasa.ac.at, WP-96-132, November 1996.
<http://www.iiasa.ac.at/Publications/Documents/WP-96-132>

Appendix

From above, DESDS model is formulated by multiobjective nonlinear programming (MONLP). MONLP has a several methods for dealing with multiobjective problems. Moreover, we can notice that the coefficients of all objective function haven't the same magnitude. For that, we tried to bring all objective functions coefficients into the same order magnitude. Firstly, the normalization concept (L_1 -norm) will be used to unite magnitude (measure unit) for all objective functions coefficients [11]. Normalization will convert the coefficients into percentages where

$$L_1\text{-norm} = \frac{v}{\|v\|_1}$$

$$\text{and } \|v\|_1 = \left[\sum_{i=1}^n |v_i| \right]$$

Now, we can select anyone of multiobjective methods. Weighting Method (WM) will be used to deal with the model [11]. In general, weights are assigned to the objective functions according to importance of each objective function of decision maker (DM). For this reason, the weights of objective functions in DESDS model will be assumed to be equal for the three objectives, where $w_1 = w_2 = w_3 = 1/3$.

The optimal solution will be obtained by solving the following weighting problem formulated:

$$\begin{aligned} &\text{Minimize} && w f(x) = \sum_{i=1}^I w_i f_i(x) \\ &\text{Subject to} && x \in X, \end{aligned} \quad (1)$$

For the purpose of certain from the goodness of DESDS model, the model will be resolved three times. Each time, each one of objective functions will be assigned the weight equal to zero. This means that this objective function will take the worst value or the maximum value of it. After that, these three values of objective functions are compared with the actual value of EOEP (see, table-13). DESDS model always achieve the results better than the results of EOEP in all cases.

	EOEP	The worth case of cost objective function $w_1=0$	The worth case of pollution objective function $w_2=0$	The worth case of subsidy objective function $w_3=0$
Cost (000' L.E)	12934930	11629484	11629484	11579358
Subsidy (000' L.E)	4224953	4131694	4131694	4162455.82
Pollution (000' PTE)	108814.4	97039.12	97039.12	95963.73

Table-13 shows the worth cases of the objective functions of DESDS model

After applying WM, the model became with one objective function. Now, we can select any language programming for solving this problem. GAMS [2] (the acronym stands for General Algebraic Modeling System) will be used as language programming for solving the weighting formulation (1). Although GAMS is optimization language programming, it hasn't a solver for dealing with multiobjective problems. GAMS is designed to make the construction and solution of large and complex mathematical programming models.

From previous words, model parameters have constant values. Furthermore, completed DESS model will be introduced again with all its values of parameters.

Minimize [225 DS_{LPG} + 246.9 $DS_{gasoline}$ + 234.94 $DS_{kerosene}$ + 225.987 DS_{coal} + 226.4 DS_{gasoil} + 176.904 $DS_{fuel\ oil}$ + 150.651 $DS_{natural\ gas}$] + [593.74 * TEP * $ES_{industry}$ + 834.004 * TEP * $ES_{residential\ commercial}$ + 512.75 * TEP * $ES_{agriculture}$ + 100.895 * TEP * $ES_{government\ public}$]

Minimize [2.9837 DS_{LPG} + 3.1046 $DS_{gasoline}$ + 3.216 $DS_{kerosene}$ + 2.9473 DS_{coal} + 3.2093 DS_{gasoil} + 3.2094 $DS_{fuel\ oil}$ + 2.6115 $DS_{natural\ gas}$]

Minimize [200 DS_{LPG} + 201.2 $DS_{gasoline}$ + 199.1 $DS_{kerosene}$ + 200.91 DS_{coal} + 191.12 DS_{gasoil} + 182.2 $DS_{fuel\ oil}$ + 100.3 $DS_{natural\ gas}$ + 90 DS_{wind} + 91.5 $DS_{hydroelectric}$ + 100.3 $DS_{natural\ gas}$]

Subject to

0.9 DS_{LPG} $TS_{LPG, industry}$ + 0.9 $DS_{gasoline}$ $TS_{gasoline, industry}$ + 0.9 $DS_{kerosene}$ $TS_{kerosene, industry}$ + 0.7 DS_{coal} $TS_{coal, industry}$ + 0.9 DS_{gasoil} $TS_{gasoil, industry}$ + 0.9 $DS_{fuel\ oil}$ $TS_{fuel\ oil, industry}$ + 0.9 $DS_{natural\ gas}$ $TS_{natural\ gas, industry}$ $\geq$ 11622658

$$0.9 DS_{LPG} TS_{LPG, transport} + 0.9 DS_{gasoline} TS_{gasoline, transport} + 0.9 DS_{kerosene} TS_{kerosene, transport} + 0.7 DS_{coal} TS_{coal, transport} + 0.9 DS_{gasoil} TS_{gasoil, transport} + 0.9 DS_{fuel oil} TS_{fuel oil, transport} + 0.9 DS_{natural gas} TS_{natural gas, transport} \geq 9634690$$

$$0.9 DS_{LPG} TS_{LPG, residential/commercial} + 0.9 DS_{gasoline} TS_{gasoline, residential/commercial} + 0.9 DS_{kerosene} TS_{kerosene, residential/commercial} + 0.7 DS_{coal} TS_{coal, residential/commercial} + 0.9 DS_{gasoil} TS_{gasoil, residential/commercial} + 0.9 DS_{fuel oil} TS_{fuel oil, residential/commercial} + 0.9 DS_{natural gas} TS_{natural gas, residential/commercial} \geq 4071075$$

$$0.9 DS_{LPG} TS_{LPG, agriculture} + 0.9 DS_{gasoline} TS_{gasoline, agriculture} + 0.9 DS_{kerosene} TS_{kerosene, agriculture} + 0.7 DS_{coal} TS_{coal, agriculture} + 0.9 DS_{gasoil} TS_{gasoil, agriculture} + 0.9 DS_{fuel oil} TS_{fuel oil, agriculture} + 0.9 DS_{natural gas} TS_{natural gas, agriculture} \geq 845000$$

$$TEP ES_{industry} \geq 2088000$$

$$TEP ES_{residential/commercial} \geq 2162000$$

$$TEP ES_{agriculture} \geq 212000$$

$$TEP ES_{government/public} \geq 18210074$$

$$ELS_{gasoil} \leq (1 - TS_{gasoil, industry} + TS_{gasoil, transport} + TS_{gasoil, residential/commercial} + TS_{gasoil, agriculture})$$

$$ELS_{fuel oil} \leq (1 - TS_{fuel oil, industry} + TS_{fuel oil, transport} + TS_{fuel oil, residential/commercial} + TS_{fuel oil, agriculture})$$

$$ELS_{natural gas} \leq (1 - TS_{natural gas, industry} + TS_{natural gas, transport} + TS_{natural gas, residential/commercial} + TS_{natural gas, agriculture})$$

$$[0.266 ELS_{gasoil} DS_{gasoil} + 0.266 ELS_{fuel oil} DS_{fuel oil} + 0.175 ELS_{natural gas} * DS_{natural gas}] + (0.307 DS_{(wind)} + 0.95 DS_{(hydroelectric)}) - TEP \geq 0$$

$$TS_{LPG, industry} + TS_{LPG, transport} + TS_{LPG, residential/commercial} + TS_{LPG, agriculture} \leq 1$$

$$TS_{gasoline, industry} + TS_{gasoline, transport} + TS_{gasoline, residential/commercial} + TS_{gasoline, agriculture} \leq 1$$

$$TS_{kerosene, industry} + TS_{kerosene, transport} + TS_{kerosene, residential/commercial} + TS_{kerosene, agriculture} \leq 1$$

$$TS_{coal, industry} + TS_{coal, transport} + TS_{coal, residential/commercial} + TS_{coal, agriculture} \leq 1$$

$$TS_{gasoil, industry} + TS_{gasoil, transport} + TS_{gasoil, residential/commercial} + TS_{gasoil, agriculture} \leq 1$$

$$TS_{fuel oil, industry} + TS_{fuel oil, transport} + TS_{fuel oil, residential/commercial} + TS_{fuel oil, agriculture} \leq 1$$

$$TS_{natural gas, industry} + TS_{natural gas, transport} + TS_{natural gas, residential/commercial} + TS_{natural gas, agriculture} \leq 1$$

$$ES_{industry} + ES_{residential/commercial} + ES_{agriculture} + ES_{government/public} \leq 1$$

$$DS_{LPG} \leq 2613625$$

$$DS_{gasoline} \leq 2500580$$

$$DS_{kerosene} \leq 1436800$$

$$DS_{coal} \leq 45000$$

$$DS_{gasoil} \leq 8000848$$

$$DS_{fuel oil} \leq 6261404$$

$$DS_{natural gas} \leq 19390000$$

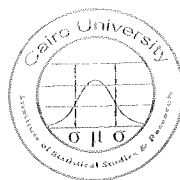
$$DS_{wind} \leq 12000$$

$$DS_{hydroelectric} \leq 2997000$$

$$DS_i, TEP, TS_{i,j}, ES_j, ELS_i \geq 0$$



The 42nd Annual Conference on Statistics,
Computer Sciences and Operation Research
2-5 Dec,2007



Outputs of GAMS Program: The output from GAMS is produced on one file, which is designed to be read by human. It can of course be printed, erased or archived like any other file. The output file is the only place where GAMS can write results for inspection or for use by other programs. Figure-9 shows the final report of GAMS output file. To void the large domain of problem, 000'L.E / 000'petroleum ton is used for simplification as the measure unit of model variables.

GAMS Rev 121 Windows NT/95/98
(PESS model)
Execution

07/18/06 22:21:39 PAGE :

```

---- 200 VARIABLE DS.L domestic supply
LPG      1864.625, gasoline 2500.580, kerosene 1536.800
coal     39.000,  gasoil  8000.848, fueloil  6661.404
naturalgas 19048.019, wind 12.000, hydroelectric 5997.000

---- 200 VARIABLE TS.L thermal share
          industry transport residential agricultu-
LPG      0.941
gasoline          1.000
kerosene          0.634
coal            1.000
gasoil           0.155
fueloil          0.900
naturalgas      0.584
              0.232

---- 200 VARIABLE ES.L electric share
industry      0.255, residentialcommercial 0.264
agriculture   0.259, governmentpublicutilities 0.222

---- 200 VARIABLE ELS.L
gasoil 0.737, fueloil 0.105, naturalgas 0.220

---- 200 VARIABLE TEP.L
VARIABLE z.L      = 8131.067 electric sector
                   = 4380.508 value of objective
                           function
    
```

Figure-9 shows the final report of GAMS output file

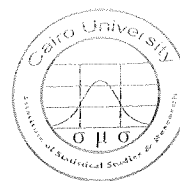


Cairo University

Institute of Statistical Studies and Research

**The 42nd Annual Conference on Statistics,
Computer Sciences and Operation Research
2-5 Dec, 2007**





Exploring Parallel Optimization Using Probability Collectives

Wael Ratch

Mohammed A. El-Beltagy

Mohamed H. Rasmy

Abstract

Traditional optimization techniques rely on a centralized approach for obtaining in their pursuit of optimality. An alternate approach, Probability Collectives (PC) theory provides a new framework for performing parallel optimization. Instead searching for variables that minimizes the objective function, PC finds a probability distribution over the variables that optimizes an associated expected value of the objective function. In this paper, we provide an overview of PC theory and detail an algorithm for performing optimization derived from it. The approach is demonstrated on two discrete optimization problems, one of them is unconstrained and the other one is constrained.

Introduction

The connections between bounded rational game theory, information theory, statistical mechanics, and Product Distribution (PD) theory provide a new framework for performing distributed optimization. The fields of statistical physics, game theory, and distributed control/optimization share one fundamental characteristic: they are all concerned with how the probability distribution governing a distributed system is related to the objective function. This shared characteristic provides the basis for translating the concepts and theory between these fields and creating large common mathematical framework that connects them. Inspired by building such bridges and recognizing their impact on optimization, the theory of Probability Collectives (PC) was developed [1, 2]. PC shares certain features with the two popular metaheuristic approaches of Simulated Annealing (SA) [3] and Genetic Algorithms (GA) [4]. The three approaches search randomly for an optimal solution and try to balance between exploring new solutions and exploiting the good solutions. Like SA, PC probability distribution to sample the new solution. However, unlike SA the search is conducted over the space of probability distributions not directly over the variables themselves, PC theory states that "instead searching for variables that minimizes the objective function, it finds a probability distribution over the variables that optimizes an associated expected value of the objective function" [5]. PC is somewhat similar to GA in that the optimization is distributed over a large number of agents, yet the PC agents are somewhat more complex entities compared to GA's individuals in a population.

Formulating the problem as a distributed stochastic optimization allows for the Product Distributions (PD) to be used in large array of problems in machine learning, statistics, multi-agent systems, and game theory. In addition, approaches can be developed, in those problem areas, which take advantage of the local structure of certain optimization problems. PD theory is quite flexible in that it handles constraints. The core insight of PD theory is to concentrate on how the agents update the probability distributions across their possible actions rather than specifically on the joint action generated by sampling those distributions [6]. First, we review PC and how to use it to formalize bounded rationality. We then review two of the previously explored techniques for finding bounded rational equilibrium, Brouwer updating and Nearest Newton updating. The version of Probability Collectives (PC) considered here, involving product distributions, is called "Product Distribution" (PD) theory [2, 7, 8, 9, 6].

1 Game Theory and Product Distribution (PD)

Probability collective (PC) theory can be considered as an extension to conventional game theory [10]. In any non-cooperative game, we have N of agents, each agent η independently selects its strategy z_η at any point of time. If an agent plays only one strategy across the game, we say the player is using pure strategy. Instead, if the agent samples its probability distribution, we say that player is using mixed strategy, i.e. matching pennies game. Let each agent have a probability distribution $q_\eta(z_\eta)$ over its strategies, i.e. a mixed strategy of its pure strategy z_η . Each player also has its own cost function $g_\eta(z)$, which is the negative of its utility function; this function is parametrized by the collective behavior of all agents in the game [11], where the goal is the minimization of objective functions rather than maximization, after each agent chooses its action, the joint-actions for all players is z and the joint probability distribution governing the joint-action of all players is a product distribution of all the N players' probability distribution, since all agents take their actions or moves independently, the product distribution form is:

$$P(z) = \prod_{\eta} q_{\eta}(z_{\eta}) \quad (1)$$

Given the mixed strategies, each agent strives to minimize its expected cost function. The expected cost function of agent η is hence:

$$E(g_{\eta}) = \int \prod_j q_j(z_j) g_j(z) dz \quad (2)$$

The expected utility is the integration over all values of z , the integral is converted to a summation when the action space is discrete. The agent η seeks to make this integration as low as possible. The conventional game theory assumes the agents are fully rational and can compute any thing related to the game. In this sense, given the strategies of the other agents, $z_{(\eta)}$, agent η with full rationality searches for z_{η}^* that satisfies the following formula:

$$z_{\eta}^* = \arg \min_{z_{\eta}} g_{\eta}(z_{\eta}, z_{(\eta)}), \quad (3)$$

where $z_{(\eta)}$ is a the strategies of all agents other than η .

According to the previous equation, z_{η}^* is the best response of agent η given the actions of the others. Every player adopts the mixed strategy that maximizes its expected utility, given the mixed strategies of the other players. Conventional non-cooperative game theory assumes each player η is "fully rational" i.e., he is able to solve for that optimal q_{η} , and he uses that optimal distribution. The joint-moves, z , is Nash equilibrium if for all players, no agent can change its mixed strategy and in the same time improves its cost function given the actions of the other agents are fixed [10]. In the real world, this assumption of full rationality almost never holds, whether the players are humans, animals, or computational entities. Even if the players are able to find the optimal mixed strategies, the cost of computation (money or time) is still an obstacle to achieve that task optimally. Because of the unrealistic assumption limits the application of game theory in real world problems. [12, 13, 7]. There are several major problems with the concept of Nash equilibrium that are still outstanding [2].

2 Combining Game Theory and Statistical Physics

A prior knowledge is required to infer the probability distribution over the states of the system. This knowledge consists of the expectation values of certain functions associated with the system's states. In statistical physics, if the system's expected energy is known this means that the probability distribution p governing the system is the Boltzmann distribution. Intuitively, players in a

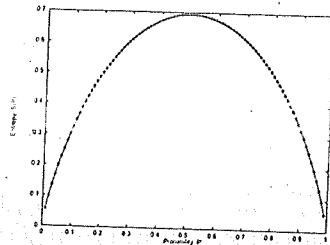


Figure 1: Entropy of probability distribution with 2 states. This graph peaked at 0.5.

game and their reward functions are mathematically identical to particles in a system and their energy functions. Games with varying numbers of agents are analogous to physical systems with varying numbers of particles. Quantified bounded rationality of agents is formally identical to the temperature of particles. This analogy opens up possibilities where the techniques of game theory and statistical physics can be combined and exploited (e.g., for discrete, continuous, mixed, and constrained optimization). By this identity the rich methods and techniques that used in statistical physics, that has been proved their applicability, can be exploited in conventional game theory to accommodate bounded rationality.

Wolpert shows that [2] in statistical physics, full rationality corresponds to a statistical mechanics system with zero level temperature, in this case the joint actions of all agents is a Nash equilibrium. Bounded rationality corresponds to non-zero temperature. In bounded rationality, the equilibrium is determined given the information available to all players. This amount of information can be quantified by the negative of Shannon entropy [14, 15] as follows:

$$S(P) = - \int P(z) \ln(P(z)) dz \quad (4)$$

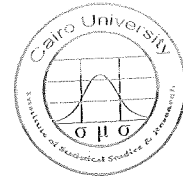
$S(P)$ measures the amount of information that can be inferred from a distribution P . $S(P)$ reaches its maximum in the case of a uniform distribution; where all possible outcomes have the same probability of occurring. So, the uniform distribution is a distribution with minimal information. The uniform distribution is the most uncertain or uninformative distribution. For example consider discrete states $z = \{z_1, z_2, \dots, z_N\}$ and assume all the states are equally likely, $P(z_\eta) = 1/N \quad \forall \eta$, then the entropy of the probability distribution $P(z)$ is:

$$\begin{aligned} S(P) &= - \sum_{\eta} P_{\eta} \ln(P_{\eta}) \\ &= - \sum_{\eta} \frac{1}{N} \ln\left(\frac{1}{N}\right) = \ln(N) \end{aligned} \quad (5)$$

By contrast, the most informative distribution is the one has a probability one at a specific value of z and all other values have probability of zero, and in this distribution the amount of uncertainty equals to 0. Figure 1 shows the entropy of probability distribution over 2D variable space.

3 The Maximum Entropy (Maxent) Lagrangian

Given a situation where we have limited observations about moves made in particular game, and we want to infer the probability that some strategy will be followed in the game, we apply the



concept of maximum entropy [14, 15]. The Maxent principle states: "The probability distribution that uniquely represents or encodes our state of that information is the one that maximize the uncertainty measure (Entropy) while remaining consistent with our information" [14]. In applying maxent we search for the probability distribution that has a maximum entropy (uncertainty) or the less informative distribution possible to be consistent with the given information. Choosing a distribution with the maximum uncertainty rather than the one with lower uncertainty, since the distribution with minimum entropy would violate the constraints of the information we do possess. So choosing one with maximum entropy is the only reasonable distribution to capture all the information given.

Formally, for simplicity assumes a finite number of possible moves (pure strategies) $|z_\eta|$ for each player η . Also, consider a common payoff (team game), in which all players share the same cost function, $G(x)$. The optimization problem can be written as:

$$\begin{aligned} \min_q E_q(G) &= \int G(z) \prod_j q_j(z) dz, \\ \text{subject to:} \quad & S(q) = s, \\ & \int q_i(z_i) dz_i = 1, \\ & q_i(z_i) \geq 0 \end{aligned} \quad (6)$$

The first constraint is the entropy constraint. At a particular level of uncertainty s , we want to find a joint probability distribution q that minimizes the utility function G . The second and third constraints are set to restrict the unit simplex and non-negative probabilities. We can convert this constrained optimization problem to equivalent unconstrained optimization problem by using the Lagrangian:

$$\begin{aligned} L(q) &= E_q(G) - T(S(q) - s) \\ &= \int G(z) \prod_j q_j(z) dz - T(S(q) - s), \end{aligned} \quad (7)$$

where T is the Lagrangian multiplier (referred to as temperature in statistical physics [2]).

To find the minimum of the original problem, we need to find the Lagrangian's saddle point. We set the partial derivatives of L with respect to T and q to zero, the Karush-Kuhn-Tucker conditions, $\partial L / \partial T = \partial L / \partial q = 0$ [16]. Solving the Lagrangian function result in a set of coupled Boltzmann equations [17, 18, 19, 1] one for each agent i :

$$q_i(z_i) \propto e^{-\frac{1}{T} E_{q_{(-i)}}(G|z_i)} \quad (8)$$

The subscript $q_{(-i)}^T$ on the expectation value indicates that the expectation is calculated according to probability distribution $\prod_{j \neq i} q_j$ and T is set to enforce entropy constraint. From equation 8, the probability of player i choosing a pure strategy z_i depends on the effect of that choice on the utilities of the other players. Trying to solve such a couple system of equations can prove quite difficult. We seek to find a formulation of the optimization problem, where we can focus only on the behavior of player i . This can be expressed as the following lagrangian minimization:



$$\begin{aligned}
 \min_{q_i} L_i(q_i) &= E(G) - TS(q_i) \\
 &= \int q_i(z_i) E(G|z_i) dz_i - TS(q_i) \\
 &= \int q_i(z_i) \prod_j q_j G(z_{(i)}, z_i) dz_i + T \int q_i(z_i) \ln q_i(z_i) dz_i, \\
 \text{s.t. :} \quad &\int q_i(z_i) dz_i = 1, \quad q_i(z_i) \geq 0 \quad \forall z_i,
 \end{aligned} \tag{9}$$

where agent i tries to minimize this local Lagrange function assuming the reward function of agent i is a team game, which is world utility of the collective. The first term in the local Lagrange function is minimized by fully rational player. By setting temperature parameter T equals to zero, this means that the agent works theoretically according to game theory and the collective strategies of the agents is a Nash equilibrium. The second term in the local Lagrange is minimized by an irrational player. This term minimized by uniform mixed strategies, i.e. for non-zero temperature. Therefore, the temperature parameter T is trades-off between rational and irrational behavior. Also we can see temperature T as a weight that balances between the expected utility and the uncertainty associated with the probability distribution across all i 's actions. As $T \rightarrow \infty$ the agent selects the irrational actions, explore new solutions in its actions space. As time increases we gradually decrease the temperature to exploit the good solutions, the agent selects the full rational actions at $T \rightarrow 0$. It is not preferable and desirable to reduce temperature to zero [20, 17, 21].

4 Optimization of the maxent Lagrangian

In this section we will explore the techniques used to minimize the Lagrangian function. Wolpert et al. [11, 19, 1, 2] introduced three basic techniques that can be used to minimize $L(q, T)$ for fixed T . the first one is based on the first order descent which is gradient descent, second technique is based on the second order descent using Hessian matrix which is Nearest Newton descent and the last one is based on solving set of coupled Boltzmann equations. For all three techniques that we will discuss we assume discrete variable space and number of agents is fixed.

4.1 First Order Gradient Descent

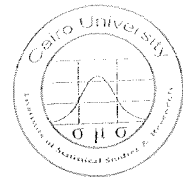
The gradient descent finds the direction that result in maximum drop in the Lagrangian $L(q, T)$. The gradient is calculated for agent action pair. The gradient descent for agent i at specific action j is

$$q_i(j) \leftarrow q_i(j) + \alpha \delta q_i(j), \tag{10}$$

$$\delta q_i(j) = u_i(j) - \sum_{z_i} \frac{u_i(z_i)}{|z_i|} \tag{11}$$

$$u_i(j) = E(G|z_i = j) + T \left[\ln \left(q_i(z_i = j) \right) \right] \tag{12}$$

Where α is the learning rate or step size and $\delta q_i(j)$ is the direction that make the large drop in $L(q, T)$ which is used by each agent i to update its probability distribution for each action j . we



subtract the second term, average change over all agent actions, to keep the distribution q to be in Q , where Q is the space or the universe of probability distribution.

So, there is one update equation for each agent action pair. Therefore, each agent has a set of update equations, one for each action. The number of update equation for agent i equals to number of actions that it can perform, $|z_i|$. These update equations are fully distributed among agents, in the sense that each agent's update at time t is independent of any other agents' update at that time. In other words, each player is responsible for updating its mixed strategies (probability distribution) and independent of the others. Each agent knows its probability distribution q_i and hence it knows $\ln q_i$, but it doesn't know G and $q_{(i)}$. So, it can't evaluate $E(G|z_i = j)$ in closed form. We can overcome this problem by using Monte Carlo sampling. The agents sample distribution q to produce the set joint actions z and the associated global utility. After the sampling each agent has a set of $(z, G(z))$ pairs. From these samples the agents are able to estimate $E(G|z_i = j)$ by using average or regression [2].

4.2 Second Order Newton Descent

This update method depends on the second derivative of the Lagrangian function with respect to $q_i(j)$. Newton update starts by making a quadratic approximation of over the space of q to the maxent Lagrangian $L(q, T)$. The Newton update rule searches for the optimal q that minimize the quadratic approximation of the Lagrangian. The benefit from quadratic approximation is that the maxent Lagrangian becomes convex [2, 12]. Searching the optimal joint probability distribution from initial distribution is equivalent to find a distribution p that has the minimal Kullback-Liebler (KL) distance [2]. The Newton update rule can be written as the follows:

$$q_i(j) \leftarrow q_i(j) - \alpha q_i(j) X \left[E(G|z_i = j) - E(G) + T \left(S(q_i) + \ln(q_i(j)) \right) \right] \quad (13)$$

Where α is the learning rate or the step size that can move in the direction of minimization of the maxent Lagrangian. The step size can be fixed over all the solution or can vary with time. In the early steps of simulation the step size is sufficiently large to enable the agents to learn. As time increases, we decrease temperature slowly until it reaches to very small value. Note that learning rate is the same for all agents [11] and does not change with time. At each iteration, we must ensure that all probabilities are non-negatives and they sums to one for each agent. Using this update rule, there is no guarantee that the sum of probabilities is one and all probabilities is non-negative. The solution: for each negative probability set its value to a very small positive value, 10^{-6} and then renormalize the p associated probability distribution. As the gradient descent, Newton update is also distributed but $E(G|z_i = j)$ can't be evaluated in closed form and we can overcome it by using Monte-Carlo sampling. The derivation of this update rule can be found in [11]. Because the Newton update rule is based on quadratic approximation of the maxent Lagrangian, it is proved a good and rapid convergence in many problems [22, 20, 23, 24].

4.3 Brouwer Update Rule

Another technique used to update the probability distribution q is based on Boltzmann equation discussed before. For discrete space agent i can update its probability distribution using the following equation:

$$q_i(j) = \frac{e^{-\frac{1}{T} E(G|z_i=j)}}{\sum_{z'_i} e^{-\frac{1}{T} E(G|z'_i=j)}} \quad (14)$$



This update rule doesn't take into consideration the previous values of probabilities and the entropy associated to its probability distribution. Brouwer update rule needs only $E(G|Z_i = j)$ that can be evaluated by using Monte-Carlo sampling. This technique is also called parallel Brouwer update since all agents update their probability distributions simultaneously; it has no guarantee to converge [25]. Another version of update rule that may converge to optimal is the serial Brouwer update where only one agent updates its probability distribution at a time. The selection of an agent for serial update can be done randomly or based on specific criteria which makes the largest drop in the maxent Lagrangian [26, 12, 13]. Another variant of Brouwer update rule is the mixed parallel-serial Brouwer update.

5 Private Utility function

Selecting the private utility function of agents has a great effect on the convergence. PC theory focused on how the agents update their probability distributions across all possible actions. All updating rules we introduced assume that all agents receive a private utility as the world utility, team game. Team game is factored; has zero Bias (aligned) with respect to objective function (world utility) but it has high variance; has poor learnability, the agents can't discern the effect of their actions on their private utility which makes the learnability process difficult. The cost associated with implementing team game is low, since it needs only one evaluation for each Monte-Carlo sample.

In COIN theory, Wolpert *et al.* introduced a utility function called the Wonderful Life Utility (*WLU*). *WLU* is proven theoretically and experimentally to have high factoredness and is also highly learnable. Using team game requires global communication among agents. Also team game performs poorly when there are a communication restriction among agents. Tumer and Agogino [27] introduced variants of *WLU* that perform nicely when there are communication restrictions among agents.

6 PC Theory For Distributed Optimization

Gradient based-optimization fails to solve non-smooth and/or discontinuous functions. It has a serious disadvantage which is the convergence to a local minima, especially if the objective is a multi-modal non-convex function. Also when the problem is integer or mixed integer, the search in variable space is a very difficult task. Although, genetic algorithm (GA) and simulated annealing (SA) use probabilities, this probability used to trade-off exploration and exploitation. But, these methods are still searching in the variable space. In PC theory, rather than searching the optimal solution in the variable space, PC theory transforms the problem from variable space ζ to probability distribution space Q . PC theory assigns for each variable (agent) in the optimization problem a probability distribution q_i over its possible values, ζ_i . By this way, PC theory searches for the joint-probability distribution q , which minimizes the associated expected objective. This transformation from searching in variable space to probability distribution space, enables us to use the gradient descent methods in non-smooth functions, integer and mixed-integer problem.

Consider the following optimization problem:

$$\begin{aligned}
 & \min_{\vec{z}} C(\vec{z}) \\
 & \text{s.t. :} \\
 & \Phi_j(\vec{z}) \leq C_j \quad \forall j = 1, 2, \dots, k, \\
 & \vec{z} \in \vec{\zeta}.
 \end{aligned} \tag{15}$$

where $\vec{z}$ is a N dimensional vector of discrete variables and $\Phi_j(\vec{z}) \leq C_j$ is set of k constraints. Before we formalize PC maxent Lagrangian we should convert this constrained problem to an equivalent un-constrained problem using Lagrange multipliers. The new unconstrained problem is:

$$G(\vec{z}) \leftarrow G(\vec{z}) + \sum_{j=1}^k \lambda_j (C_j - \Phi_j(\vec{z})) \quad (16)$$

Where λ_j is the Lagrange multipliers and $(C_j - \Phi_j(\vec{z}))$ are non-negative. To find the update function for each Lagrangian multiplier λ_j , take the derivative of the augmented Lagrangian with respect to each λ_j as follows:

$$\frac{\partial}{\partial \lambda_j} \left(G(\vec{z}) + \sum_{j=1}^k \lambda_j (C_j - \Phi_j(\vec{z})) \right) = (C_j - \Phi_j(\vec{z})) \quad (17)$$

Then the update rule of λ_j is:

$$\lambda_j \leftarrow \lambda_j + \mu (C_j - \Phi_j(\vec{z})) \quad (18)$$

Where μ is the step size, it may be fixed for all constraints and may vary across time. The expectation is taken over all Monte-Carlo samples. This form of un-constrained problem can be solved by GA or SA. To solve this optimization problem by PC we should transform this problem from searching in variable space ζ to searching in probability distribution space Q . This transformation is done by constructing the maxent Lagrangian. Firstly, solving the optimization problem as a collective, we should represent each variable z_i as an agent, i , in a collective. Let q_i be the probability distribution over all possible values of z_i . All agents take their actions independently, so the joint probability distribution of the collective is the product distribution of all agents probability distribution q_i , $q(z) = \prod_j q_j(z_j)$. Because of the team game's poor perormace in a large scale collective, we will assign agents a Wonderful life utility as a private utility, $g_i(\vec{z}) = WLU_i(\vec{z}) = G(\vec{z}) - G(Z_{(i)})$, the Clamped term set to null. The maxent Lagrangian problem of each agent is given below:

$$\begin{aligned} \min_{q_i} L_i(q_i) &= \int q_i(z_i) \prod_j q_j g_i(z_i, z_{(i)}) dz_i \\ &\quad + T \int q_i(z_i) \ln(q_i(z_i)) \\ \text{s.t. :} \quad &\int q_i(z_i) dz_i = 1, q_i(z_i) \geq 0 \forall z_i. \end{aligned} \quad (19)$$

7 Solution Algorithm

As we discussed before this maxent Lagrangian equations are distributed among agents. The algorithm of solving this distributed optimization is given in figure 2.

The algorithm consists of three main stages: Monte-Carlo sampling using their probability distributions, private function evaluation for each agent and probability distribution update of agents. The steps of the algorithm are:

1. Intialization:

- Set parameters: temperature T , learning rate to update probability distributions α , constraints update step size μ for constrained optimization only.

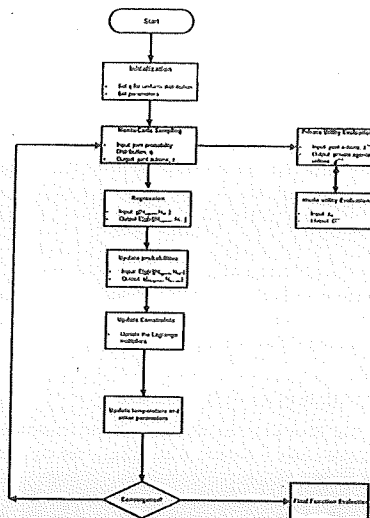


Figure 2: Solution algorithm flow chart

- Set the annealing schedule function as a function of the number of iterations, i.e. $T_t = T_0 0.99^t$ where t is the current iteration and T_0 is the initial temperature.
- Set the number of Monte-Carlo samples N_{MC} ; N_{MC} should be sufficient to sample all space of z and it has great effect on the performance and convergence of the algorithm.
- Set the starting probability distribution of each agent (variable) to uniform distribution over its action space.
- Set the termination criteria. There are many criteria can be used as a stopping condition such as: probability distributions convergence, $\|\sum_i q_i^t - q_i^{t-1}\|$ where $\|\cdot\|$ is the Euclidean distance, and the objective function convergence is $|G^t(z) - G^{t-1}(z)|$.

2. Minimizing the maxent Lagrangian:

- Monte-Carlo samples the agent probabilities to obtain N_{MC} IID joint action samples; z is a matrix of size $N \cdot N_{MC}$.
- For each sample z^m :
 - (a) Evaluate the objective function.
 - (b) For each agent, evaluate its private utility based on the Wonderful life utility. Note, this step is eliminated if we set the private utilities of the agents as a team game.
- Compute the expected utilities of agents given each possible action j . For discrete variable we use a simple average as follows:

$$E(g_i | z_i = j) = \frac{\sum_{N_{MC}} g_i(z_i, z_{(i)}) \Theta(z_i = j)}{\sum_{N_{MC}} \Theta(z_i = j)}, \quad (20)$$

where $\Theta(Z_i = j)$ is an indicator function that equals to 1 when the action selected is j and 0 otherwise. We can use data aging to reuse information from the previous samples. The data aging controlled by aging parameter γ and the equation of the conditional expectation with data aging for discrete variables is:

$$E(g_i | z_i = j) = \frac{N_{ij}^t}{D_{ij}^t} = \frac{\sum_{N_{MC}} g_i(z_i, z_{(i)}) \Theta(z_i = j) + \gamma N_{ij}^{t-1}}{\sum_{N_{MC}} \Theta(z_i = j) + \gamma D_{ij}^{t-1}}, \quad (21)$$

where N_{ij}^t and D_{ij}^t are used capture the expected reward of agent i for carrying out action j at time step t . Also the data aging is useful for small samples. Note, in the first iteration we should be sure that all actions selected at least one to ensure D_{ij} doesn't include zero.

- Update the probability distributions of agents according to any one of the update rules which were discussed before, either by order gradient descent using equation 12, second order descent Newton using equation 13, or Brouwer updating using equation 14. We must ensure that the probabilities are non-negative and sums to one for each agent probability distribution.
- Update the Lagrange multipliers for each constraint.
- Update temperature using annealing equation. It is recommended to update Lagrange multiplier and temperature every l iterations to ensure the convergence of the probabilities, l can be set to 10 or 25.
- Evaluate the stopping criteria. If not satisfied go to step 2, else go to step 3.

3. Final evaluation

- Determine the highest probability action for each agent.
- Evaluate the objective function associated to this joint action.

8 Experimental Results

In this section we will introduce some experiments to explore PC theory and show it's effectiveness to solve the optimization problems.

8.1 Simple Game Experiment

A payoff table of a simple game between two bounded rational players is shown in figure 3. Each player has two actions or moves with common payoff. Each player tries to select the action that minimize the associated cost without co-operating or talking with the other player, in terms of game theory terminology they play non-cooperative game. This game has two global minima, one when player one selects action D and player two selects action L , $\{D, L\}$, and another at $\{U, R\}$.

Player one and player two are trying to minimize their local lagrangian function in equation 9. Figure 4 is a visualization of the maxent Lagrangian of player one at high and low temperature. The surface of the Lagrangian function at high temperature shows that there are no significant differences between actions for both players, the minimum of the Lagrangian when they play with a mixed strategy $(0.5, 0.5)$. As the temperature decays the surface of the Lagrangian function tends to have only one minimum as shown in the left panel of figure 4 and the game equilibrium tends to be a pure strategy in which player one plays with D and player two plays with L . The visualization in figure 4 shows that it is equally likely to the agents to select any action from their action space at high temperature, as the temperature decays the ability of players to discern the quality of

		Player 1	
		L	R
Player 2	U	0	-1
	D	-2	0

Figure 3: Simple two players game

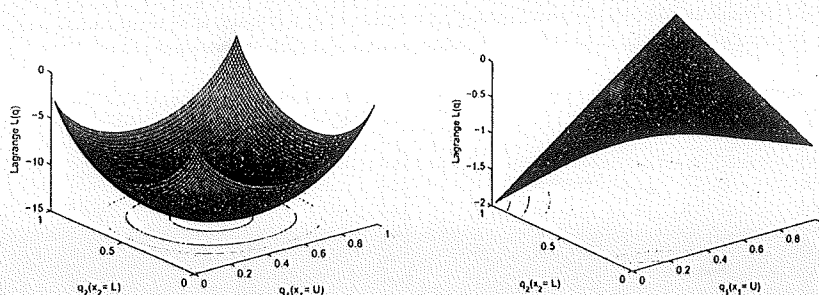


Figure 4: Visualization of the maxent Lagrangian function. The left graph is at temperature = 10 (high temperature) and the right figure is at temperature = 0.05 (low temperature)

their actions increases and they can easily differentiate among actions with local or global minimum.

The left graph of figure 5 shows the contours of the maxent Lagrangian function against the probabilities of selecting actions for both players at temperature equals to 1. The global minimum of the Lagrangian function is at the upper left corner and the local minimum in the opposite direction, lower right corner.

The right graph of figure 5 shows the solution trajectory using nearest Newton update rule. The learning rate used for these experimentations was 0.5, this value chosen to achieve good performance. In figure 6, the left graph shows the trajectory of both parallel and nearest Newton update rules. All figures plotted over the contours at temperature = 0.1. the right graph shows the solution trajectories of both parallel Brouwer and serial Brouwer update rules. In the parallel Brouwer, both players updates their probabilities in each iteration, however, for the serial Brouwer update only one agent updates its probability distribution over its action space at a time. The parallel update rule is not stable and can fall in local minima and oscillates between two minima.

8.2 Single dimension 0-1 Knapsack problem Experiments

One dimensional knapsack problem is one of the most famous problems in operations research in which there are objects of N types are available. Each article of type i has weight w_i and value v_i . A knapsack that can only hold at most a total weight of W to be loaded with these objects so as to maximize the value of the objects included subject to the weight capacity. The object takes value of zero if it is not included in the Knapsack and a value of one if it is included in the

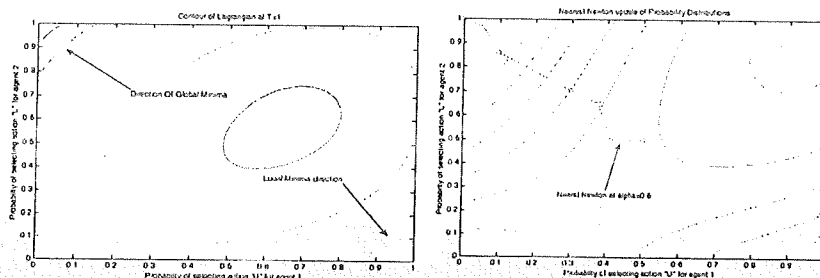


Figure 5: Visualization of the maxent Lagrangian function. The left graph is at temperature =10 (high temperature) and the right figure is at temperature =0.05 (low temperature)

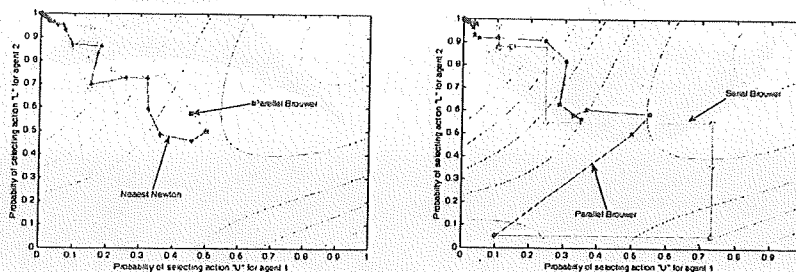


Figure 6: Visualization of the maxent Lagrangian function. The left graph is at temperature =10 (high temperature) and the right graph is at temperature =0.05 (low temperature)

knapsack. 0 – 1 knapsack has many applications in capital budgeting, project selection,...etc. Let x_i be a binary variable that takes value one if it is included and zero otherwise. Formally, x_i can be defined as follows:

$$x_i = \begin{cases} 1 & \text{if article } i \text{ is selected} \\ 0 & \text{Otherwise} \end{cases} \quad (22)$$

The problem can be formulated as follows:

$$\begin{aligned} \max_{x_i} Z(x) &= \sum_i v_i x_i \\ \text{s.t.} \quad & \sum_i w_i x_i \leq W, \\ & x_i = 0 \text{ or } 1 \quad \forall \quad i = 1, 2, \dots, N. \end{aligned} \quad (23)$$

In this experiment, we will use simple application of knapsack problem, project or investment selection problem. The data is given in table 1, we have seven investments. The board needs to select subset of these projects that maximizes the total return without exceeding the total available budget. The world utility of the system can be defined as following:

Investment(x_i)	Cost(w_i)	Value(v_i)
1	3	12
2	4	12
3	3	9
4	3	15
5	15	90
6	13	26
7	16	112
Weight(W)	35	

Table 1: : data for the investment selection problem

$$G(x) = -Z(x) = -\sum_i v_i x_i \quad (24)$$

The Knapsack problem is a constrained problem, so we must handle the constraint by augmenting this constraint into the objective function with Lagrangian multiplier λ as follows:

$$G(x) = -\sum_i v_i x_i + \lambda \left(W - \sum_i w_i x_i \right) \quad (25)$$

In this experiment we used two types of private utilities, Team Game and Wonderful life utility, as a signal to indicate the quality of the taken action. In team Game utility each agent receives the world utility (with augmentation) as its private reward, whereas in Wonderful life utility, each agent receives the following reward as a signal:

$$WLU_i(x) = \begin{cases} -v_i - \lambda w_i & \text{if } x_i = 1 \\ 0 & \text{Otherwise} \end{cases} \quad (26)$$

We treat each variable as an agent. This agent has a probability distribution over its action space which includes only two actions, select investment or not. We used only 10 Monte-Carlo samples per iteration. This problem has only one global minimum at $x_4 = x_5 = x_7 = 1$.

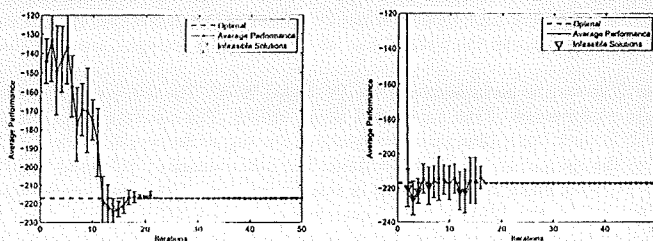


Figure 7: The convergence of the knapsack problem based on team game utility. The right graph is the average performance of the augmented objective function. The left graph is the convergence of the objective function at the maximum probabilities.

Figure 7 shows the convergence of the performance of the probability collective theory based on team game utility, it shows the convergence of the average performance of the world utility with

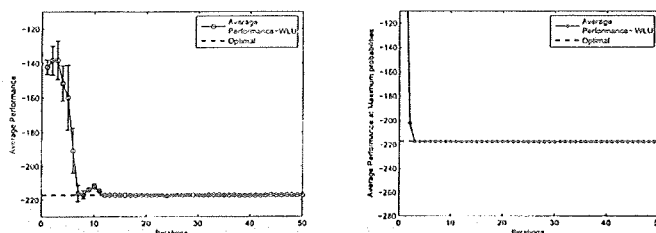


Figure 8: The performance of the knapsack problem based on wonderful life utility. The right graph is the average performance of the augmented objective function. The left graph is the convergence of the objective function at the maximum probabilities.

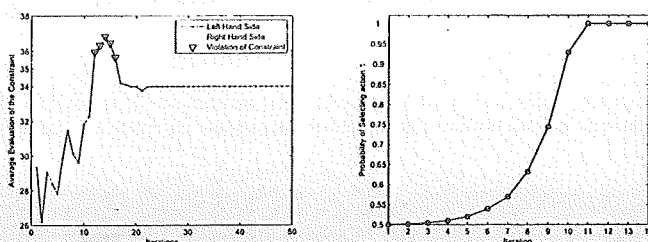
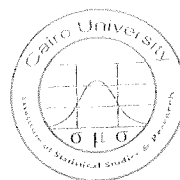


Figure 9: The left chart is the average evaluation of the knapsack capacity constraint. The right chart is The evolution the Probability of selecting action 0 for the first agent, x_1 .

time. Startly, the agents try to take actions which minimize their cost function. But, they have no experience about these actions so they need to learn. During the learning process the agents take actions that make their cost function low as possible which may lead to infeasible joint actions (solution). This clear in the figure, all values marked by a triangle are infeasible although they are better than optimal. It reached to optimal solution after 26 iterations (left graph). The right graph shows the convergence of the objective to optimal. The evaluation was at the maximum probabilities not the average. The performance of the PC-Based-TG characterized by high variance until it converges. All values under the optimal line which marked by triangles are infeasible.

Figure 8 shows the performance of the PC-based-WLU. The left graph shows convergence of the average performance of the augmented world utility. The convergence of the algorithm PC-based-WLU is much faster than the convergence of the augmented world utility in PC-based-TG, it converge to the optimal on the ninth iteration. The performance of PC-based-WLU is faster than PC-based-TG by a factor of about 300% in this experiment. This result confirms the theoretical reasons that tell us the WLU is much more learnable than TG especially when the system is congested with agents. Also, the right graph shows the convergence of the objective function at the maximum probabilities. The algorithm reaches the optimal in the third iteration; this is much less than the convergence on the PC-based-TG. Also the PC-based-WLU is characterized by having very low variance compared with PC-based-TG since the clamped term removed much of the noise caused by the other agents' actions. This makes the decision process of the agents more clear and



effective, because each agent discern the impact of its action on the world utility.

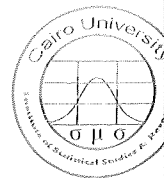
The left graph of figure 9 shows the value of the capacity constraint with time. The constraint is violated for all solutions which make the constraint value under the dashed line (maximum capacity). The right graph of figure 9 shows the evolution of the probability of selecting action 1 for the first agent (variable) across time. This figure shows how the agent learnt to exploit the action which minimize its cost function.

Conclusion and future work

From these experiments, the probability collective theory showed that it is a strong and a promising approach to distributed or parallel optimization. We aimed to demonstrate the veracity of PC on a classical operations research problem and the results are encouraging that we seek to apply it to other problems. Also, the robust mathematical base of the probability collective theory makes it a superior approach to handle many types of optimization problems, especially combinatorial problems that makes the search space more complex. An added advantage is that it can be used to explore the robustness of the optimal solution by exploiting resultant joint probabilities of actions.

References

- [1] Chiu Fan Lee and David H. Wolpert. Product distribution theory for control of multi-agent systems. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems*, New York, NY, 2004.
- [2] David Wolpert. Information theory - the bridge connecting bounded rational game theory and statistical physics, 2004/Feb/19/ 2004. TY - ELEC.
- [3] S. Kirkpatrick Vecchi, C. D. Gelatt, and M. P. Optimization by simulated annealing. *Science*, Number 4598, 13 May 1983, 220, 4598:671-680, 1983.
- [4] David E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Professional, January 1989.
- [5] Stefan Bieniański, David H. Wolpert, and Ilan Kroo. Discrete, continuous, and constrained optimization using collectives, 2004.
- [6] W. Macready, S. Bieniański, and D. Wolpert. Adaptive multi-agent systems for constrained optimization. 2004.
- [7] David H. Wolpert. Bounded rationality and statistical physics. In Akira Namatame et al, editor, *In Proceeding of WEHIA*. Springer-Verlag, in print, 2004.
- [8] Kristina Lerman, Aram Galstyan, Kagan Tumer, and David Wolpert. Collectives and design of complex systems. pages 231-256. Springer, 2004.
- [9] William Macready and David H. Wolpert. Distributed optimization. In *Presented at the International Conference on Complex Systems*, Boston, MA, 2004.
- [10] H. Scott Bierman and Luis Fernandez. *Game Theory with Economic Applications*. Addison - Wesley, 2 edition, 1998.
- [11] Stefan Bieniański. *Distributed Optimization and Flight Control Using Collectives*. PhD thesis, Stanford University, 2005.



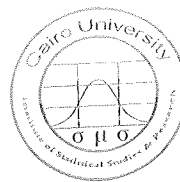
- [12] David H. Wolpert. Finding bounded rational equilibria, part i : Interactive focusing. In *Proceeding of the Eleventh International Symposium on Dynamic Games and Applications*, Tucson Arizona, 2004.
- [13] D.H. Wolpert. Finding bounded rational equilibria part ii: alternative lagrangians and uncountable move spaces, 2004.
- [14] David Mackay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.
- [15] Thomas Cover and Joy Thomas. *Elements of Information Theory*. Wiley-Interscience, 1991.
- [16] Jonathan Borwein and Adrian S. Lewis. *Convex Analysis and Nonlinear Optimization: Theory and Examples*. Springer-Verlag, second edition, 2005.
- [17] John W. Lawson Wolpert and David H. Adaptive programming of unconventional nano-architectures. *Journal of Computational and Theoretical Nanoscience*, 3:272, 2006.
- [18] David H. Wolpert and Chiu Fan Lee. An adaptive metropolis-hastings scheme: sampling and optimization. *Europhysics Letters*, 76:353, 2006.
- [19] D. H. Wolpert and S. Bieniawski. Theory of distributed control using product distributions. In *in Proceedings of CDC04*, 2004.
- [20] Stefan R. Bieniawski, Ilan M. Krooy, and David H. Wolpert. Flight control with distributed effectors. 2005.
- [21] D. H. Wolpert. "adaptive distributed control:beyond single-instant categorical variables,". In A. S. et al, editor, *in Proceedings of MSRAS04*. Springer Verlag, 2004.
- [22] Chien-Feng Huang, Stefan Bieniawski, Charlie E. M. Strauss, and David H. Wolpert. A comparative study of probability collectives based multiagent systems and genetic algorithms. In *Proceedings of the 2005 conference on Genetic and evolutionary computation*, pages 751 – 752, Washington DC, USA, 2005. ACM Press New York, NY, USA.
- [23] Geoff Ryder and Kevin Ross. A probability collectives approach to weighted clustering algorithms for ad hoc networks, October 26, 2005 2005.
- [24] Nicolas E. Antoine, Stefan Bieniawski, David H. Wolpert, and Ilan Kroo. Fleet assignment using collective intelligence, 2004.
- [25] Stefan Bieniawski and David H. Wolpert. Product distributions for distributed optimization, 2004.
- [26] David Wolpert. A predictive theory of games, 2005/Dec/8/ 2005. TY - ELEC.
- [27] A. Agogino and K. Tumer. Team formation and communication restrictions in collectives. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multi-Agent Systems*. Melbourne, Australia, 2003.



Cairo University

Institute of Statistical Studies and Research

**The 42nd Annual Conference on Statistics,
Computer Sciences and Operation Research
2-5 Dec, 2007**

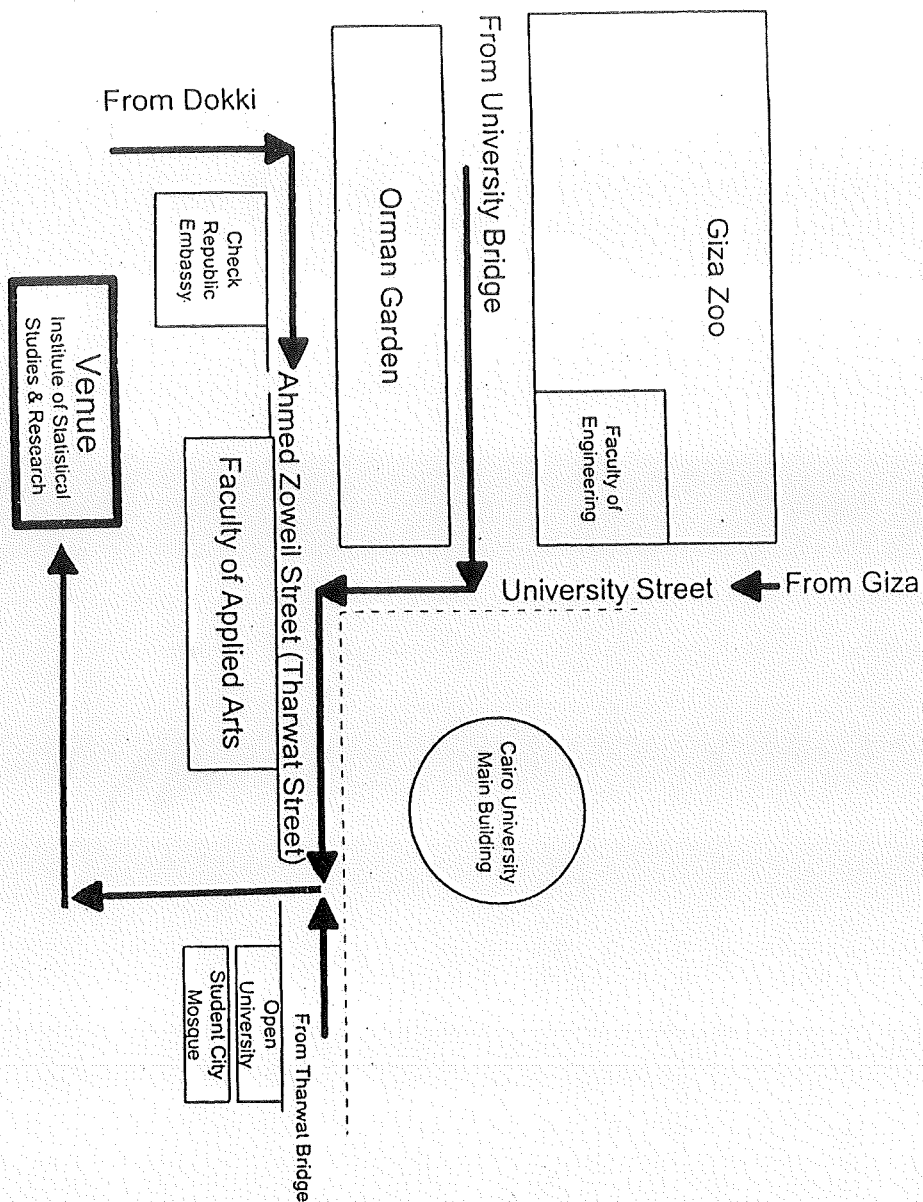




Cairo University

Institute of Statistical Studies and Research

The 42nd Annual Conference on Statistics,
Computer Sciences and Operation Research
2-5 Dec, 2007





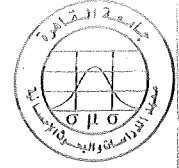
جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
للإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



الجهاز المركزي للتعبئة العامة والإحصاء	مديحة عبد الله محمد
الجهاز المركزي للتعبئة العامة والإحصاء	مريم إبراهيم ضلالم
وزارة الزراعة	مصطفى بدير
وزارة التجارة والصناعة	مصطفى عبد الغفار
كلية الحاسبات - جامعة القاهرة	معتز خورشيد
وزارة الزراعة	ممدوح فياض
طالب بحث - معهد الإحصاء - جامعة القاهرة	مني كيلاني
طالب بحث	مهند سعد عبد كركوش
الولايات المتحدة الأمريكية	نادجيب بوزار
وزارة الزراعة	نادي فوزي
الجهاز المركزي للتعبئة العامة والإحصاء	ناهد محمد عبد الحميد
وزارة الزراعة	نبيل زكي عبد الصمد
كلية العلوم - جامعة سوهاج	نوميري علام
وزارة التجارة والصناعة	نيفين حسين
الولايات المتحدة الأمريكية	هدى عبد العاطي زهدي
كلية الحاسبات - جامعة القاهرة	هدي محمد انسي
الجهاز المركزي للتعبئة العامة والإحصاء	هناء عبد العظيم خالد
جامعة حسينية - الجزائر	هني محمد نبيل
كلية التجارة - جامعة حلوان	هويدا محمد عز الدين
معهد الإحصاء - جامعة القاهرة	وائل راتب محمود
أكاديمية السادات للعلوم الإدارية	وليد محمد عفيفي
جامعة الإمارات العربية	يحيى البسيوني
وزارة الزراعة	يحيى سليمان



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الإثني والأربعون
للإحصاء وعلماء الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



كلية الحاسبات - جامعة القاهرة	محمد حسن رسمي
وزارة التجارة والصناعة	محمد حسن شحاتة
وزارة التجارة والصناعة	محمد حسين سعد الدين
مركز البحوث الزراعية - وزارة الزراعة	محمد رافع
وزارة الزراعة	محمد رشاد
أكاديمية الإسكندرية للإدارة والمحاسبة	محمد رضا صبحي محمد
كلية التجارة - جامعة بنها	محمد زايد الدسوقي
الهيئة القومية للأفناق	محمد سامي محمد
طالب بحث - معهد الإحصاء - جامعة القاهرة	محمد سعد محمد
كلية الحاسبات - جامعة القاهرة	محمد عادل البلتاجي
طالب بحث - معهد الإحصاء - جامعة القاهرة	محمد عبد الله
معهد الدراسات والبحوث البيئية - جامعة المنوفية	محمد عبد الله عباس عطوه
طالب بحث - كلية التجارة - جامعة الإسكندرية	محمد علي عبد العزيز غريب
معهد الإحصاء - جامعة القاهرة	محمد عويس
طالب بحث - معهد الإحصاء - جامعة القاهرة	محمد فتحي السنطاوي
طالب بحث - معهد الإحصاء - جامعة القاهرة	محمد محمود السيد
معهد الإحصاء - جامعة القاهرة	محمد مصطفى حسن
كلية الحاسبات - جامعة القاهرة	محمد مصطفى صالح
وزارة التجارة والصناعة	محمد نبيل الشيمي
معهد الإحصاء - جامعة القاهرة	محمد نجيب
جامعة مؤتة - الأردن	محمود الحيلان
معهد الإحصاء - جامعة القاهرة	محمود رياض
الهيئة المصرية للرقابة على التأمين	محمود عبد العزيز السخاوي
كلية التربية - الرياض - السعودية	محمود هندي
وزارة التجارة والصناعة	مختار الشريف



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



كلية الحاسبات-جامعة القاهرة	علي علي فهمي
معهد الإحصاء - جامعة القاهرة	علي وجدي
جامعة الكويت	عماد الدين علي
مركز البحوث الزراعية - وزارة الزراعة	عماد سداد عبد الحكيم
المجلس الأعلى للأثار - وزارة الثقافة	عماد مقلد
كلية العلوم - جامعة جنوب الوادي	عمرو رضوان
كلية الحاسبات-جامعة القاهرة	عمرو كامل
كلية العلوم - جامعة المنوفية	عمرو مسعد صابر
كلية الحاسبات - جامعة القاهرة	عبد محمد عمري
الجهاز المركزي للتعبئة العامة والإحصاء	فاطمة محمد العشري
الشركة العامة للكهرباء	فرج إبراهيم عاشور
وزارة الزراعة	فرغلي جاهين
كلية التجارة - جامعة بنها	فريد جرجس
كلية الحاسبات - جامعة القاهرة	كريم درويش
كلية العلوم - جامعة القاهرة	ليلى عبد العال
وزارة التجارة والصناعة	ماجدة الحسيني سليط
كلية التجارة- جامعة الأسكندرية - فرع دمنهور	ماجدة محمد حجاج
الجهاز المركزي للتعبئة العامة والإحصاء	ماجدة محمد عبد الحميد
وزارة التنمية الإدارية	مجدى الحناوي
الأكاديمية العربية للعلوم والتكنولوجيا والنقل البحري	مجدى عبد المنعم
وزارة التجارة والصناعة	محسن الخضيرى
كلية التجارة - جامعة بنها	محمد إبراهيم خليل
وزارة التجارة والصناعة	محمد إبراهيم محمد
طالب بحث - معهد الإحصاء - جامعة القاهرة	محمد أبو بكر إبراهيم
طالب بحث	محمد السيد الداودي



جامعة القاهرة

معهد الدراسات والبحوث الإحصائية

المؤتمر السنوي

الثاني والأربعون

للإحصاء وعلمو الحاسب وبحوث العمليات

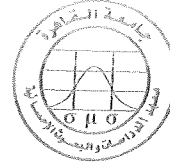
٢-٥ ديسمبر ٢٠٠٧



وزارة التجارة والصناعة	شريف أحمد شريف
الهيئة العامة لشئون الزراعة والثروة السمكية - الكويت	صدفة الكندري
معهد الإحصاء - جامعة القاهرة	صلاح مهدي
طالب بحث - معهد الإحصاء - جامعة القاهرة	طاهر طه زكي
طالب دكتوراه - كلية الحاسبات - جامعة القاهرة	طلال سعيد صالح
جامعة مؤتة - الأردن	طلال علي الهواري
كلية الحاسبات - جامعة القاهرة	عاصم ثروت
أكاديمية البحث العلمي - أسوان	عاطف حسنين خالد
كلية العلوم - جامعة القاهرة	عبد الجبار علي التميمي
الهيئة المصرية للرقابة على التأمين	عبد الحكيم توفيق
مركز البحوث الزراعية - وزارة الزراعة	عبد الرحمن السيد محمد
طالب بحث - معهد الإحصاء - جامعة القاهرة	عبد الرحيم عبد الصبور
كلية التجارة - جامعة بنها	عبد الفتاح قنديل
كلية العلوم - جامعة المنصورة	عبد الفتاح مصطفى
وزارة التجارة والصناعة	عبد النبي عبد المطلب
طالب بحث - معهد الإحصاء - جامعة القاهرة	عبير بحبي محمد
كلية الحاسبات - جامعة القاهرة	عثمان حجازي
الجهاز المركزي للتعبئة العامة والإحصاء	عزة محمد عبده
كلية التجارة - جامعة حلوان	عصام أبو القاسم
الهيئة المصرية للرقابة على التأمين	عصام عبد الغني صبره
معهد الإحصاء - جامعة القاهرة	علاء احمد عبد العزيز
كلية الحاسبات - جامعة مصر للعلوم والتكنولوجيا	علاء زغول محمود
كلية الحاسبات - جامعة القاهرة	علي البسطاويسي
كلية إدارة الأعمال - جامعة الملك سعود - السعودية	علي بن محمد الجمعه
كلية العلوم الإدارية - جامعة الكويت	علي صالح يوسف



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



وزارة التجارة والصناعة	حسين عمران
كلية العلوم الإدارية - جامعة الكويت	حسيني ابراهيم حمدي
طالب بحث - جامعة سوهاج	حمدي حسن السيد
كلية الحاسبات - جامعة القاهرة	خالد مصطفى السيد
كلية الحاسبات - جامعة القاهرة	دعاء صالح علي
طالب بحث - سلطنة عمان	رابعة محمد عبد الله اللويهي
كلية التجارة - جامعة بنها	رأفت شاكور
كلية العلوم - جامعة طنطا	رشا فتح الله العجمي
وزارة الزراعة	رفعت حنفي سالم
وزارة الزراعة	رمزي محمد مبارك
وزارة الزراعة	رمضان علي عثمان
مدارس نفرتاري الدولية - القاهرة	رمضان معوض
كلية الحاسبات-جامعة القاهرة	ريم بهجت
كلية الحاسبات - جامعة المنيا	سارة حسن ابراهيم
كلية العلوم - جامعة قاريونس - ليبيا	سالم احمد بن عمران
القوات المسلحة	سامي محمد أحمد عطية
كلية التجارة - جامعة بنها	سحر رأفت
كلية التربية- جامعة للحديدة - اليمن	سلطان علي محمد سالم
كلية الحاسبات-جامعة القاهرة	سمحاء البلتاجي
معهد الإحصاء - جامعة القاهرة	سمير عاشور
كلية الحاسبات-جامعة القاهرة	سمير عبد الرحمن
كلية الحاسبات-جامعة القاهرة	سناء العلا حنفي
الولايات المتحدة الأمريكية	سناء طرمان
الجهاز المركزي للتعبئة العامة والإحصاء	سناء محمد عباده
وزارة التجارة والصناعة	سهير أبو العنين



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
للإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



وزارة التنمية الإدارية	أشرف عطية
وزارة الزراعة	الجبالي أبو السعود
طالب بحث - جامعة الطائف - المملكة العربية السعودية	الجوهرة الجعيد
معهد الإحصاء - جامعة القاهرة	السيد أحمد الشربيني
المركز القومي للبحوث	السيد أحمد متولي
معهد الإحصاء - جامعة القاهرة	السيد خاطر
كلية الحقوق - جامعة المنصورة	السيد عبد الخالق
مدارس نفرتاري الدولية - القاهرة	أمانى محمد أنور
الجهاز المركزي للتعبئة العامة والإحصاء	أميرة جمال الدين
الجهاز المركزي للتعبئة العامة والإحصاء	آيات محمد عبد العاطي
كلية التجارة - جامعة الأزهر	إيمان أحمد الدسوقي
شركة التعاون للبترو	إيمان أحمد الشرقاوي
الجامعة العمالية	أيمن علي موسى
جامعة حسبية - الجزائر	بلعوز بن علي
وزارة الزراعة	بيومي عبد المجيد
كلية العلوم بنات - جامعة الأزهر	تغريد عبد الرحمن
جامعة الإمارات العربية	توفيق زبيدي
معهد الإحصاء - جامعة القاهرة	توفيق يونس
الجهاز المركزي للتعبئة العامة والإحصاء	ثروت فايق نخلة
وزارة الزراعة	جمال عبد الباقي عبد الهادي
جامعة بنسلفانيا - الولايات المتحدة الأمريكية	جوليان ملبين
طالب بحث - المملكة العربية السعودية	الجوهرة الجعيد
كلية التجارة - جامعة بنها	حازم سرور
معهد الإحصاء - جامعة القاهرة	حجازي زاهر
الهيئة القومية للأنفاق	حسين حسن فايق



كلية العلوم الرياضية - جامعة الخرطوم
المعهد العالي للدراسات التكنولوجية المتخصصة - القاهرة
كلية التجارة - جامعة حلوان
الجهاز المركزي للتعبئة العامة والإحصاء
كلية علوم وهندسة الحاسوب - اليمن
كلية الحاسبات-جامعة القاهرة
الجهاز المركزي للتعبئة العامة والإحصاء
معهد الإحصاء - جامعة القاهرة
كلية الحاسبات-جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
الجامعة الأمريكية
كلية العلوم - جامعة المنيا
كلية الهندسة - جامعة القاهرة
مركز البحوث الزراعية
المشروع العربي لصحة الأسرة
كلية التجارة - جامعة حلوان
معهد الإحصاء - جامعة القاهرة
الجهاز المركزي للتعبئة العامة والإحصاء
وزارة التجارة والصناعة
جامعة تنز - اليمن
كلية الحاسبات-جامعة القاهرة
المعمل المركزي للنظم الخبيرة
كلية التربية - الرياض - السعودية

إبراهيم العبد إبراهيم العبد
إبراهيم بغدادى
إبراهيم حسن ابراهيم
إبراهيم حلمي عبد الله
إبراهيم عبد الرب احمد
إبراهيم فرج
أبو بكر الجندي
أحمد أمين الشيخ
أحمد حسين كمال
أحمد حمزة محمد أحمد
أحمد رافع
أحمد رضوان
أحمد صوفي أبو طالب
أحمد عبد العزيز عطية
أحمد عبد المنعم
أحمد عبد الهادي
أحمد فؤاد عطية
أحمد كمال عبد العزيز
أحمد كمال فريد
أحمد محمد غالب الرباصي
أدهم شاهر
أسامة إسماعيل
أسماء عايض الراددي



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثنائي والأربعون
الإحصاء وعلوم الحاسب وبيوتات العمليات
٢-٥ ديسمبر ٢٠٠٧



وحدة التصوير والطباعة والتجليد

- أ. غريب محمود سالم
- أ. أحمد عبد اللطيف محمد
- أ. السيد محمد سعد
- أ. يوسف لطفي محمد
- أ. محمد سعيد سيد
- أ. رضا كمال حسن

الإستعلامات

- أ. صلاح عمر



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



اللجنة الفنية للمؤتمر

- د. أحمد أمين الشيخ
 - أ. ياسمين محمد إبراهيم
 - أ. نادية حلمي عواد
- سكرتارية المؤتمر

مساعدا اللجنة الفنية للمؤتمر

- أ. جليلة كامل
- أ. سوسن عبد الحليم
- أ. سلمى محمود أحمد
- أ. صفاء محمود
- أ. هالة اسماعيل
- أ. نسرین إمام متولي
- أ. سمير بونان
- أ. مصطفى عباس
- أ. منى زغلول



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



- د. مرفت غيث
أ.د. محمد جمال الدين درويش
أ.د. محمد حسن رسمي
د. محمد مصطفى حسن
أ.د. محمد نجيب عبد الفتاح
أ.د. محمود رياض محمود
أ.د. معتز خورشيد
أ.د. منى خليفة
د. مها مصطفى كمال
د. ليلي ناصف
د. ناجى درويش رمضان
د. نجلاء عبد المنعم مراد
أ.د. نيفين درويش
د. هشام نبويه المهدي
د. هشام احمد حفنى
أ.د. هشام حسن مخلوف
د. وليد محمد عرفة
أ.د. يوسف محبوب
- معهد الإحصاء - جامعة القاهرة
عميد كلية الحاسبات - جامعة الاهرام الكندية
كلية الحاسبات - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
رئيس قسم بمعهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
نائب رئيس اتحاد الجامعات العربية
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
كلية الهندسة - جامعة القاهرة
كلية الحاسبات - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



- أ.د. الهام شكري محمد
د. توفيق محمد يونس
أ.د. حجازي محمد زاهر
أ.د. دعد محمد فؤاد
أ.د. ريم بهجت رضا
د. رمضان عبد الحميد زين الدين
د. سحر أحمد نيازي
د. سرحان حسين سرحان
أ.د. سعد الدين محمد الشيبال
أ.د. سمير كامل عاشور
د. سمية العطار
د. سمية محمود السعدني
أ.د. سناء العلا حنفى احمد
د. سهام اسماعيل ميرة
أ.د. سيد مشعال
د. صلاح مهدى محمد
أ.د. عاطف محمد خليفة
أ.د. عاطف محمد عبد المنعم
أ.د. عبد الغنى محمد عبد الغنى
د. عبد الله عبد الفتاح
أ.د. عثمان حجازي محمد
د. علاء أحمد عبد العزيز
أ.د. علي علي فهمي
أ.د. غزال عبد العزيز عامر
أ.د. فاروق تمام على شعيب
د. فاطمة الليثي
أ.د. ليلى عبد العزيز
أ.د. مجدى سيد مصطفى
- عميد معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
وكيل كلية الحاسبات - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
وكيل كلية الحاسبات - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
رئيس قسم - معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
مدير المركز الديموجرافي - القاهرة
رئيس قسم - معهد الإحصاء - جامعة القاهرة
كلية الحاسبات - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
عميد كلية الحاسبات - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة
معهد الإحصاء - جامعة القاهرة



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلموه الحاسبي وبحث العمليات
٢-٥ ديسمبر ٢٠٠٧



أمانة المؤتمر والهيئة العلمية

رعاية المؤتمر

يعقد المؤتمر تحت رعاية الأستاذ الدكتور رئيس الجامعة أ.د. علي عبد الرحمن يوسف

رئاسة المؤتمر

رئيس المؤتمر
عميد معهد الدراسات والبحوث الإحصائية
أ.د. إلهام شكري محمد

أمانة المؤتمر

أمين عام المؤتمر
وكيل معهد الدراسات والبحوث الإحصائية
لشئون خدمة المجتمع وتنمية البيئة
أ.د. أماني موسى محمد

الهيئة العلمية مرتبة أبجدياً

الإسم	الجهة
أ.د. أحمد البدوي عبادة سرحان	معهد الإحصاء - جامعة القاهرة
د. أحمد أمين الشيخ	معهد الإحصاء - جامعة القاهرة
أ.د. أحمد فؤاد محمد عطية	وكيل بمعهد الإحصاء - جامعة القاهرة
أ.د. أحمد حسن يوسف	معهد الإحصاء - جامعة القاهرة
أ.د. أحلام النجار	معهد بحوث الاقتصاد الزراعي
أ.د. الحسيني عبد البر راضى	معهد الإحصاء - جامعة القاهرة
أ.د. أماني موسى محمد موسى	وكيل بمعهد الإحصاء - جامعة القاهرة



• تنمية القدرات التصديرية للمشروعات الصغيرة والمتوسطة

بالاشتراك مع وزارة التجارة والصناعة.

• تجربة محافظة المنوفية الناجحة في تنفيذ برامج تنظيم الأسرة.

كما تتناول جلسات المؤتمر وعددها خمسة عشر جلسة تشمل ما يزيد عن الستين بحثا .

إن من أسباب نجاح المؤتمر السنوي للمعهد واستمراره واستقراره وتطوره ترجع الى عدة عوامل من أهمها :

- الدعم الدائم لإدارة جامعة القاهرة لهذا المؤتمر ومساندته علي مدى سنوات انعقاده .

- العدد الكبير من المشاركين فيه ببحوثهم العلمية المتميزة من داخل مصر وخارجها.

- مشاركة الجهات المختلفة من وزارات وأجهزة الدولة والقطاعين العام والخاص الى جانب بعض الهيئات الدولية والمحلية من خلال تنظيم الندوات وتقديم الدعم لمؤتمرنا هذا.

ولا يسعني في النهاية الا الترحيب بجميع المشاركين في هذا المؤتمر من داخل مصر وخارجها مع التمنيات بدوام التقدم لمصرنا الحبيبة.

عميد المعهد

أ.د. الهام شكرى محمد سعيد



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



- كلية الإعلام جامعة القاهرة
 - اتحاد الجامعات العربية
 - جامعة القاهرة فرع الخرطوم
 - كلية الحاسبات والمعلومات
- وكلنا أبناء هذا المعهد نفخر بأن نكون جميعا من نبت هذا الأصل الطيب والقيم
الرفيعة لأساتذتنا الأجلاء.

وعلي محور آخر يحرص دائما مؤتمرنا السنوي علي تقديم العديد من الندوات
ومن بينها:

- أهم مؤشرات النتائج الأولية للتعداد العام للسكان والإسكان
والمنشآت - ٢٠٠٦
- بناء قاعدة بيانات الأسر المصرية.
- واقع المرأة المصرية.
- تحسين الإحصاءات الزراعية باستخدام أساليب المعاينة
الإحصائية بالاشتراك مع وزارة الزراعة وإستصلاح الأراضي.
- وسائل تنمية الصادرات المصرية وذلك بالاشتراك مع وزارة
التجارة والصناعة - قطاع بحوث التسويق والدراسات السلعية
والمعلومات.
- فرص العمل المتاحة لخريجي قسم الإحصاء.



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



في العيد الستيني للمعهد نتوجه بمزيد من الشكر والعرفان لكل من أسهم خلال مسيرة المعهد وهم :

• ا.د. عبد المنعم الشافعي (صاحب الفكرة والإهتمام بتدريس الإحصاء في المجالات التطبيقية لخدمة الجهات التي تحتاج للعمل الإحصائي وتدريبها بطريقة علمية سليمة على فترة طويلة (مدة الدبلوم))

• ا.د. أحمد عباده سرحان عميد المعهد الأسبق وهو الذي يرجع إليه الفضل في تطوير المعهد من شغله جزءا من مبنى صغير إلى هذا الصرح الكبير وكذلك بناء عدد كبير من كوادر أعضاء هيئة التدريس الحاصلين على الدكتوراه من الخارج .

• ا.د. محمد مرسي أحمد - ا.د. عبد العزيز السيد
(دعما المعهد في رسالته وصدور قرارات إنشائه وتنوع دبلوماته وشهاداته في الستينات والسبعينات)

• ا.د. حسن محمد حسين - ا.د. سعد الدين الشبال واللذان أكملتا مسيرة التطور بالمعهد علميا وتكوين كوادر من أعضاء هيئة التدريس.

ويمثل خريجي المعهد وأعضاء هيئة التدريس به عصب العملية التعليمية كل في تخصصه ليس في هذا المعهد فحسب وإنما في جميع الجامعات المصرية والعربية وإحتل بعضهم مناصب رفيعة في الجامعات الأمريكية والأوروبية وكانوا الأساس الأكاديمي لجميع كليات الحاسبات بجمهورية مصر العربية.

وقد كان المعهد دائما حريصا على مد يد العون للكليات المختلفة، فقد إستضاف المعهد خلال مسيرته كل من:



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧

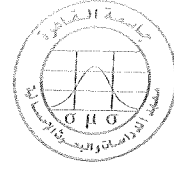


كلمة السيد الأستاذ الدكتور
إلهام شكري محمد
عميد المعهد
معهد الإحصاء ستون عاما من العطاء

يأتي إنعقاد مؤتمرنا السنوي في عامه الثاني والأربعون ليواكب التطور الهائل في علوم الإحصاء والحاسب ونظم المعلومات وبحوث العمليات ونود أن نعرض هنا نبذة تاريخية عن هذا المعهد العريق وبدايته في عام ١٩٤٧ تابعاً لكلية التجارة جامعة القاهرة . وفي عام ١٩٦٢ تم تأسيسه كمعهد علمي مستقل من معاهد جامعة القاهرة وكان يشغل جزءاً من مبنى صغير في شارع المساحة بالدقي . وفي ديسمبر ١٩٦٩ إنتقل المعهد إلى مبناه الحالي وتم إفتتاح مركز الحسابات العلمية تابعاً للمعهد وعقد مع هذا الإفتتاح مؤتمر علمي دولي كبير وإستضاف المعهد عدداً كبيراً من الباحثين من مصر والدول العربية والأوروبية والولايات المتحدة الأمريكية وقد ساهم ذلك في فتح مجالات الاتصال بالعلماء في الخارج للباحثين من المعهد وإستطاع عدد كبير من معيدي المعهد في هذا الوقت السفر على منح دراسية وبعثات للدكتوراه للدراسة بالولايات المتحدة وأوروبا بالإضافة إلى حصول عدد كبير من العاملين بوحدة الحاسب الآلي بالمعهد على منح تدريبية وكان معظمها في أوروبا حيث كان المعهد من الجهات التي تملك منذ عام ١٩٦٩ جهاز حاسب من طراز ICL 1905 والذي كان أضخم حاسب آلي في مصر في ذلك الوقت.



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧





جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون

الإحصاء وعلاوه الحاسبه وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



فهرس الإحصاء التطبيقي

٤٤-١٨	١ مفهوم الخوارزميات الجينية وتطبيقاتها في مجال الاقتصاد هنى محمد نبيل ، بلعزوز بن على
-------	--

فهرس الإحصاء السكاني

٧٠-٤٦	١ أثر العوامل الاقتصادية والاجتماعية على استخدام خدمات رعاية الأمومة في بعض الدول العربية أحمد عبد المنعم ، ماجدة محمد عبد الحميد
٩٠-٧١	٢ توقعات الخصوبة بالمناطق الجغرافية في مصر حتى عام ٢٠١٧ السيد خاطر ، محمد مصطفى ، منى كيلاني
١١٤-٩١	٣ مستويات واتجاهات العقم الاولى في مصر محمد نجيب عبد الفتاح ، محمد مصطفى حسن ، عيبر يحيى محمد
١٤٥-١١٥	٤ مستوى وخصائص المتعطلين في مصر ٢٠٠٥-٢٠٠٠ فاطمة محمد العشري ، مريم ابراهيم ضلام ، محمد حسين كسير



جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
للإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧





جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



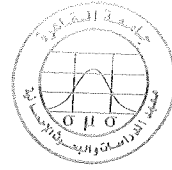
بسم الله الرحمن
الرحيم

وقل اعملوا فسيرى الله عملكم
ورسوله والمؤمنون

صدق الله العظيم

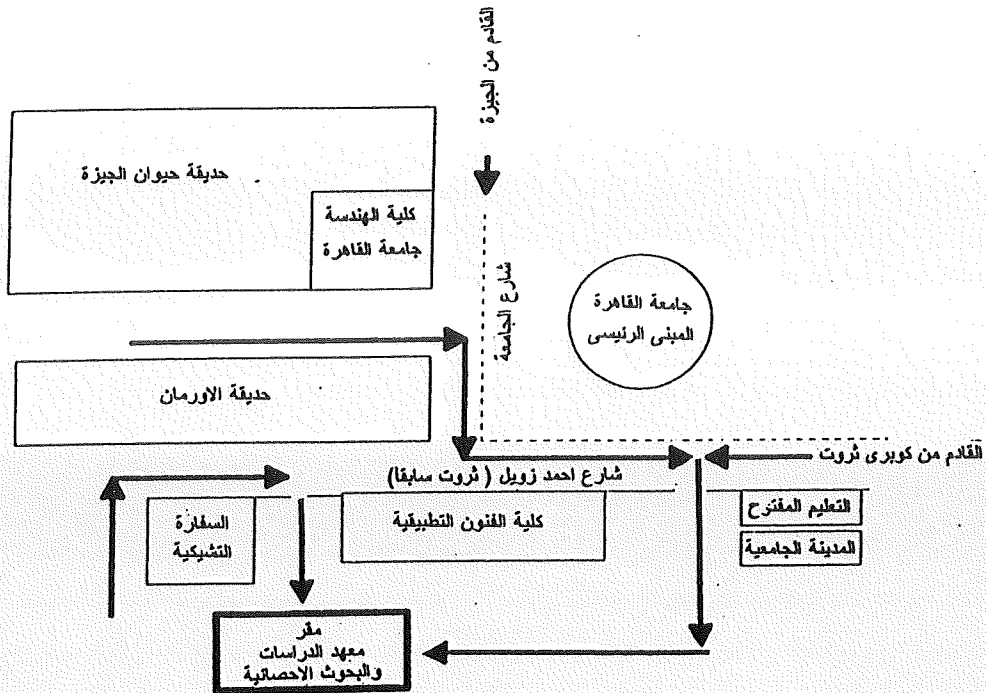


جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
الإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧





جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
النهائي والأربعون
الإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧



٥- شارع د. أحمد زويل (ثروت سابقا) - الأورمان - جيزة
جمهورية مصر العربية

تليفون: ٠٠٢٠٢٣٣٣٥٣١٦١

٠٠٢٠٢٣٣٣٥٨٣٥٥

٠٠٢٠٢٣٣٣٥٨٤٩٦

فاكس: ٠٠٢٠٢٣٧٤٨٢٥٣٣

E-mail: issr2@frcu.eun.eg

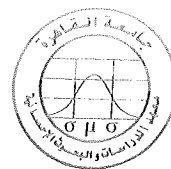
موقع المعهد على شبكة المعلومات الدولية:

www.issr.edu.eg



جامعة القاهرة

جامعة القاهرة
معهد الدراسات والبحوث الإحصائية
المؤتمر السنوي
الثاني والأربعون
للإحصاء وعلوم الحاسب وبحوث العمليات
٢-٥ ديسمبر ٢٠٠٧





مجلد

علوم الحاسب - بحوث العمليات