Studies in Computational Intelligence 302

Bijaya Ketan Panigrahi
Ajith Abraham
Swagatam Das (Eds.)

Computational Intelligence in Power Engineering

Springer

Studies in Computational Intelligence, Volume 302

Editor-in-Chief
Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Further volumes of this series can be found on our homepage: springer.com

Vol. 280. Chang Wen Chen, Zhu Li, and Shiguo Lian (Eds.)
Intelligent Multimedia Communication: Techniques and Applications, 2010
ISBN 978-3-642-11685-8

Vol. 281. Robert Babuska and Frans C.A. Groen (Eds.)
Interactive Collaborative Information Systems, 2010
ISBN 978-3-642-11687-2

Vol. 282. Husrev Taha Sencar, Sergio Velastin, Nikolaos Nikolaidis, and Shiguo Lian (Eds.)
Intelligent Multimedia Analysis for Security Applications, 2010
ISBN 978-3-642-11754-1

Vol. 283. Ngoc Thanh Nguyen, Radoslaw Katarzyniak, and Shyi-Ming Chen (Eds.)
Advances in Intelligent Information and Database Systems, 2010
ISBN 978-3-642-12089-3

Vol. 284. Juan R. González, David Alejandro Pelta, Carlos Cruz, Germán Terrazas, and Natalio Krasnogor (Eds.)
Nature Inspired Cooperative Strategies for Optimization (NICSO 2010), 2010
ISBN 978-3-642-12537-9

Vol. 285. Roberto Cipolla, Sebastiano Battiato, and Giovanni Maria Farinella (Eds.)
Computer Vision, 2010
ISBN 978-3-642-12847-9

Vol. 286. Zeev Volkovich, Alexander Bolshoy, Valery Kirzhner, and Zeev Barzily
Genome Clustering, 2010
ISBN 978-3-642-12951-3

Vol. 287. Dan Schonfeld, Caifeng Shan, Dacheng Tao, and Liang Wang (Eds.)
Video Search and Mining, 2010
ISBN 978-3-642-12899-8

Vol. 288. I-Hsien Ting, Hui-Ju Wu, Tien-Hwa Ho (Eds.)
Mining and Analyzing Social Networks, 2010
ISBN 978-3-642-13421-0

Vol. 289. Anne Håkansson, Ronald Hartung, and Ngoc Thanh Nguyen (Eds.)
Agent and Multi-agent Technology for Internet and Enterprise Systems, 2010
ISBN 978-3-642-13525-5

Vol. 290. Weiliang Xu and John Bronlund
Mastication Robots, 2010
ISBN 978-3-540-93902-3

Vol. 291. Shimon Whiteson
Adaptive Representations for Reinforcement Learning, 2010
ISBN 978-3-642-13931-4

Vol. 292. Fabrice Guillet, Gilbert Ritschard, Henri Briand, Djamel A. Zighed (Eds.)
Advances in Knowledge Discovery and Management, 2010
ISBN 978-3-642-00579-4

Vol. 293. Anthony Brabazon, Michael O'Neill, and Dietmar Maringer (Eds.)
Natural Computing in Computational Finance, 2010
ISBN 978-3-642-13949-9

Vol. 294. Manuel F.M. Barros, Jorge M.C. Guilherme, and Nuno C.G. Horta
Analog Circuits and Systems Optimization based on Evolutionary Computation Techniques, 2010
ISBN 978-3-642-12345-0

Vol. 295. Roger Lee (Ed.)
Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, 2010
ISBN 978-3-642-13264-3

Vol. 296. Roger Lee (Ed.)
Software Engineering Research, Management and Applications, 2010
ISBN 978-3-642-13272-8

Vol. 297. Tania Tronco (Ed.)
New Network Architectures, 2010
ISBN 978-3-642-13246-9

Vol. 298. Adam Wierzbicki
Trust and Fairness in Open, Distributed Systems, 2010
ISBN 978-3-642-13450-0

Vol. 299. Vassil Sgurev, Mincho Hadjiski, and Janusz Kacprzyk (Eds.)
Intelligent Systems: From Theory to Practice, 2010
ISBN 978-3-642-13427-2

Vol. 300. Baoding Liu (Ed.)
Uncertainty Theory, 2010
ISBN 978-3-642-13958-1

Vol. 301. Giuliano Armano, Marco de Gemmis, Giovanni Semeraro, and Eloisa Vargiu (Eds.)
Intelligent Information Access, 2010
ISBN 978-3-642-13999-4

Vol. 302. Bijaya Ketan Panigrahi, Ajith Abraham, and Swagatam Das (Eds.)
Computational Intelligence in Power Engineering, 2010
ISBN 978-3-642-14012-9

Bijaya Ketan Panigrahi, Ajith Abraham,
and Swagatam Das (Eds.)

Computational Intelligence in Power Engineering

Springer

Dr. Bijaya Ketan Panigrahi
Department of Electrical Engineering
Indian Institute of Technology
110016 New Delhi
India
E-mail: bijayaketan.panigrahi@gmail.com

Dr. Swagatam Das
Department of Electronics and
Telecommunication Engineering (ETCE)
Jadavpur University
Raja S. C. Mullick Road
700032 Calcutta
India

Prof. Ajith Abraham
Norwegian Center of Excellence
Center of Excellence for Quantifiable
Quality of Service
Norwegian University of Science
and Technology
O.S. Bragstads plass 2E
7491 Trondheim
Norway
E-mail: ajith.abraham@ieee.org

ISBN 978-3-642-14012-9 e-ISBN 978-3-642-14013-6

DOI 10.1007/978-3-642-14013-6

Studies in Computational Intelligence ISSN 1860-949X

Library of Congress Control Number: 2010931382

Typeset & *Cover Design:* Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

Computational Intelligence (CI) is one of the most important powerful tools for research in the diverse fields of engineering sciences ranging from traditional fields of civil, mechanical engineering to vast sections of electrical, electronics and computer engineering and above all the biological and pharmaceutical sciences. The existing field has its origin in the functioning of the human brain in processing information, recognizing pattern, learning from observations and experiments, storing and retrieving information from memory, etc. In particular, the power industry being on the verge of epoch changing due to deregulation, the power engineers require Computational intelligence tools for proper planning, operation and control of the power system. Most of the CI tools are suitably formulated as some sort of optimization or decision making problems. These CI techniques provide the power utilities with innovative solutions for efficient analysis, optimal operation and control and intelligent decision making. Due to the nonlinear, interconnected and complex nature of Power System networks and the proliferation of power electronics devices (STATCOM, UPFC, TCSC etc), the CI techniques become the promising candidates for optimal planning, intelligent operation and automatic control of the power system. Neural Network (NN), Fuzzy logic (FL) as well as the derivative free optimization techniques like Genetic Algorithm (GA), Simulated Annealing (SA) and the Swarm Intelligence (SI) techniques like Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO) play an important role in power industry for decision-making, modeling, and control problems. Due to the nonlinear nature of Power System networks and industrial electric systems (FACTS, HVDC etc), fuzzy logic and neural networks are promising candidates for planning, fault detection, automatic control, system identification, load and load/weather forecasting, etc. Distribution system routing and loss minimization are dealt with effectively using evolutionary algorithms and Swarm intelligence techniques.

This edited volume deals with different CI techniques for solving real world Power Industry problems. The technical contents will be extremely helpful for the researchers as well as the practicing engineers in the power industry.

Bjaya Ketan Panigrahi, IIT Delhi, India
Ajith Abraham, MIR Labs, USA
Swagatam Das, Jadavpur University, India

.

Contents

Robust Design of Power System Stabilizers for Multimachine Power Systems Using Differential Evolution

M. A. Abido

Abstract. In this chapter, an evolutionary algorithm based approach to multimachine Power System Stabilizer (PSS) design is presented. This approach employs Differential Evolution (DE) technique to search for optimal settings of PSS parameters. An eigenvalue-based objective function to enhance system damping for electromechanical modes is considered. The robustness of the proposed approach to the initial population is demonstrated. The performance of the proposed DE based PSS (DEPSS) under different disturbances, loading conditions, and system configurations is investigated and examined for different multimachine power systems. The eigenvalue analysis and the nonlinear simulation results show the robustness and the effectiveness of the DEPSSs to damp out the local as well as the interarea modes of oscillations and work effectively over a wide range of loading conditions.

1 Introduction

Power systems are experiencing low frequency oscillations due to disturbances. The oscillations may sustain and grow to cause system separation if no adequate damping is available. To enhance system damping, the generators are equipped with power system stabilizers (PSSs) that provide supplementary feedback stabilizing signals in the excitation systems [1-3]. DeMello and Concordia [3] presented the concepts of synchronous machine stability as affected by excitation control. They established an understanding of the stabilizing requirements for static excitation systems. In recent years, several approaches based on modern control theory have been applied to PSS design problem. These include optimal control, adaptive control, variable structure control, and intelligent control [4-8].

Despite the potential of modern control techniques with different structures, power system utilities still prefer the conventional lead-lag power system stabilizer structure [9-11]. The reasons behind that might be the ease of tuning of conventional stabilizer parameters during commissioning and the lack of assurance of the stability related to some adaptive or variable structure techniques.

M.A. Abido
Electrical Engineering Department
King Fahd University of Petroleum & Minerals
Dhahran 31261, Saudi Arabia

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 1–18.
springerlink.com

Kundur *et al* [11] have presented a comprehensive analysis of the effects of the different CPSS parameters on the overall dynamic performance of the power system. It is shown that the appropriate selection of CPSS parameters results in satisfactory performance during system upsets.

A lot of different techniques has been reported in the literature pertaining to coordinated design problem of CPSS. Different techniques of sequential design of PSSs are presented [12-13] to damp out one of the electromechanical modes at a time. However, the stabilizers designed to damp one mode can produce adverse effects in other modes. The sequential design of PSSs is avoided in [14-16] where various methods for simultaneous tuning of PSSs in multimachine power systems are proposed. Unfortunately, the proposed techniques are iterative and require heavy computation burden due to system reduction procedure. In addition, the initialization step of these algorithms is crucial and affects the final dynamic response of the controlled system. A gradient procedure for optimization of PSS parameters is presented in [17]. Unfortunately, the problem of the PSS design is a *multimodal* optimization problem, *i.e*., there exist more than one local optimum. Hence, local optimization techniques are not suitable for such a problem. Moreover, there is no local criterion to decide whether a local solution is also the global solution. Therefore, conventional optimization methods that make use of derivatives and gradients, in general, not able to locate or identify the global optimum.

Recently, heuristic search algorithms such as genetic algorithm (GA) [18-19], tabu search algorithm [20], simulated annealing [21], and particle swarm optimization (PSO) [22] have been applied to the problem of PSS design. The results are promising and confirming the potential of these algorithms for optimal PSS design. Unlike other optimization techniques, heuristic search algorithms are population-based search algorithm, which work with a population of strings that represent different potential solutions. Therefore, these algorithms have implicit parallelism that enhances their search capability and the optima can be located more quickly when applied to complex optimization problems. Unfortunately, recent research has identified some deficiencies in their performance [23]. This degradation in efficiency is apparent in applications with highly *epistatic* objective functions, *i.e*., where the parameters being optimized are highly correlated. In addition, the premature convergence of some of these algorithms degrades their performance and reduces their search capability.

More recently, a new evolutionary computation technique, called differential evolution (DE) algorithm, has been proposed and introduced [24-30]. The algorithm is inspired by biological and sociological motivations and can take care of optimality on rough, discontinuous and multi-modal surfaces. The DE has three main advantages: it can find near optimal solution regardless the initial parameter values, its convergence is fast and it uses few number of control parameters. In addition, DE is simple in coding and easy to use. It can handle integer and discrete optimization [24-30].

The performance of DE algorithm was compared to that of different heuristic techniques. It is found that, the convergence speed of DE is significantly better than that of GA [31]. In [31], the performance of DE was compared to PSO and evolutionary algorithms (EAs). The comparison was performed on a suite of 34

widely used benchmark problems. It was found that, DE is the best performing algorithm as it finds the lowest fitness value for most of the problems considered in that study. Also, DE is robust; it is able to reproduce the same results consistently over many trials, whereas the performance of PSO is far more dependent on the randomized initialization of the individuals [31]. In addition, the DE algorithm has been used to solve high-dimensional function optimization (up to 1000 dimensions) [32]. It is found that, it has superior performance on a set of widely used benchmark functions. Therefore, the DE algorithm seems to be a promising approach for engineering optimization problems. It has successfully been applied and studied to many artificial and real optimization problems [33-44].

In this study, DE based approach to PSS design is proposed. The problem of PSS design is formulated as an optimization problem with mild constraints and an eigenvalue-based objective function. Then, DE algorithm is employed to solve this optimization problem. To investigate the potential of the proposed approach, two different examples of multimachine power systems have been considered. Eigenvalue analysis and nonlinear simulation results have been carried out to assess the effectiveness of the proposed PSSs under different disturbances. In addition, the performance of the proposed DEPSS is compared to that of recent approaches reported in the literature.

2 Problem Statement

2.1 System Model and PSS Structure

A power system can be modeled by a set of nonlinear differential equations as:

$$\dot{X} = f(X, U) \tag{1}$$

where X is the vector of the state variables and U is the vector of input variables. In this study, the two-axis model [2] given in Appendix is used for nonlinear time domain simulations.

In the design of PSSs, the linearized incremental models around an equilibrium point are usually employed [1-3]. Therefore, the state equation of a power system with n machines and n_{PSS} stabilizers can be written as:

$$\Delta \dot{X} = A \Delta X + B U \tag{2}$$

where A is $4n \times 4n$ matrix and equals $\partial f / \partial X$ while B is $4n \times n_{PSS}$ matrix and equals $\partial f / \partial U$. Both A and B are evaluated at a certain operating point. ΔX is $4n \times 1$ state vector while U is $n_{PSS} \times 1$ input vector.

A widely used conventional lead-lag PSS is considered in this study. It can be described as

$$U_i = K_i \frac{sT_w}{1+sT_w} \frac{(1+sT_{1i})}{(1+sT_2)} \frac{(1+sT_{3i})}{(1+sT_4)} \Delta\omega_i \tag{3}$$

where T_w is the washout time constant, U_i is the PSS output signal at the i^{th} machine, and $\Delta\omega_i$ is the i^{th} machine speed deviation from the synchronous speed. The time constants T_w, T_2, and T_4 are usually prespecified [14]. The stabilizer gain K_i and time constants T_{1i} and T_{3i} are remained to optimize.

2.2 Objective Functions

To increase the system damping to electromechanical modes, the following eigenvalue-based objective function is considered.

$$J=min\{\zeta_I : \zeta_i \in \zeta\text{s of electromechanical modes}\} \tag{4}$$

where ζ_I is the damping ratio of the i^{th} electromechanical mode eigenvalue respectively. In the optimization process, it is aimed to *Maximize J* in order to increase the damping of electromechanical modes. The problem constraints are the optimized parameter bounds. Therefore, the design problem can be formulated as the following optimization problem.

$$\text{Maximize } J \tag{5}$$

Subject to

$$K_i^{min} \le K_i \le K_i^{max} \tag{6}$$

$$T_{1i}^{min} \le T_{1i} \le T_{1i}^{max} \tag{7}$$

$$T_{2i}^{min} \le T_{2i} \le T_{2i}^{max} \tag{8}$$

$$T_{3i}^{min} \le T_{3i} \le T_{3i}^{max} \tag{9}$$

$$T_{4i}^{min} \le T_{4i} \le T_{4i}^{max} \tag{10}$$

Typical ranges of the optimized parameters are [0.001-50] for K_i, [0.06-1.5] for T_{1i} and T_{3i}, and [0.01-0.1] for T_{2i} and T_{4i} [2]. The time constant T_w is set as 5s [19].

Considering the objective functions given in (4), the proposed approach employs DE algorithm to solve this optimization problem and search for optimal set of PSS parameters, $\{K_i, T_{1i}, T_{2i}, T_{3i}, T_{4i}, i=1,2,\ldots,n_{PSS}\}$.

3 Differential Evolution

3.1 Overview

To overcome the difficulties in traditional optimization techniques, new evolutionary population based searching techniques were proposed with promising success. In this paper, an approach based on a technique called "Differential Evolution" is presented and will be used to solve PSS design problem.

3.2 Differential Evolution Algorithm

Between the years 1994 to 1996, an optimization technique called Differential Evolution (DE) is introduced. The DE is a population based searching optimization technique and is characterized by its simplicity, robustness, few control variables and fast convergence [24-29]. Comparison between DE and the well known Particle Swarm Optimization technique in reference [30] indicated that DE has better performance. The steps of DE algorithm can be described as follows:

- **Step-1 (Initialization):**

  As any evolutionary algorithm, the DE works using set of candidate solutions forming the problem population. An initial population of *NP* candidate solutions is randomly generated over the problem space:

$$x_{ij} = x_{j\text{-lower limit}} + r * (x_{j\text{-upperr limit}} - x_{j\text{-lower limit}}) \tag{11}$$

  where r is the rondom number $r \in [0,1]$.

- **Step-2 (Objective Evaluation):**

  The objective of each solution is computed.

- **Step-3 (Stopping Criteria):**

  The stopping criteria are checked. These are the conditions under which the search process will terminate. In this study, the search will terminate if one of the following criteria is satisfied: **(a)** the number of iterations since the last change of the best solution is greater than a prespecified number; or **(b)** the number of iterations reaches the maximum allowable number.

- **Step-4 (Mutation):**

  For each vector x_i in the population, a mutant vector is created by randomly selecting three vectors (R1, R2, R3), different from x_i and different from each other.

$$x^{'} = x_{R1} + F^{*} (x_{R3} - x_{R2}) \tag{12}$$

  where, F is a mutation factor between [1,0] and it controls the solution variation in the mutation stage.

- **Step-5 (Crossover):**

  To further perturb the mutant vector, a trial vector is generated by copying parameters from the parent solution X_i and the mutant vector $X^{'}_i$ in a probabilistic manner to generate a Trail Vector $X^{''}_i$ to be used in the selection stage. This is accomplished by comparison between a randomly generated number and a specified crossover factor (CR) between [1, 0], as explained in Fig. 1. However, in case CR=0, then all parameters are copied from parent vector X_i except one, randomly selected from the mutant vector X'_i. Whereas, if CR=1 then all parameters are copied from mutant vector X'_i except one, randomly selected from the parent vector X_i.

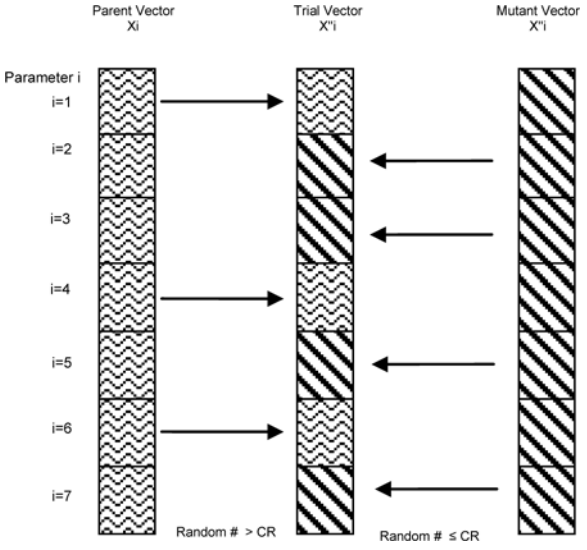


Fig. 1. Crossover Process

- **Step-6 (Selection):**

  Selection is performed using the objectiv values. To keep the population size constant over subsequent generations, the selection process is carried out to determine which one of the child and the parent will survive in the next generation. If the child yields a better value of the fitness function, it replaces its parent in the next generation; otherwise, the parent is retained in the population. Hence the population either gets better in terms of the fitness function or remains constant but never deteriorates.

3.3 DE Implementation

The proposed DE based approach was implemented using the FORTRAN language and the developed software program was executed on a 2.66-GHz PC. Initially, several runs have been done with different values of DE key parameters such as differentiation (or mutation) constant *F*, crossover constant *CR*, and size of population *NP*. In this paper, the following values are selected:

F=0.9; *CR*=0.5; *NP*=100; the search will be terminated if **(a)** the number of iterations since the last change of the best solution is greater than 100; or **(b)** the number of iterations reaches 5000.

To demonstrate the effectiveness of the proposed design approach, two different examples of multimachine power systems are considered. In both examples, PSS parameters are optimized at several operating condition to ensure the robustness of the proposed PSSs. These conditions represent different loading conditions and system configurations. It is worth mentioning that the nonlinear system model is used in time-domain simulations.

4 Example 1: Three Machine Power System

4.1 Test System and PSS Design

In this example, the 3-machine 9-bus system shown in Fig. 2 is considered. The rated MVA of G_1, G_2, and G_3 are 247.5, 192, and 128 respectively. Details of the system data are given in [1]. The participation factor method shows that the generators G_2 and G_3 are the optimum locations for installing PSSs. Hence, the optimized parameters are K_i, T_{1i}, and T_{3i}, i=2,3. The optimization process was carried out at the operating points specified as *base case, case* 1*, and case* 2. The system and generator loading levels at this case are given in Table 1 and Table 2 respectively.

To demonstrate the robustness of the proposed approach to the initial solution, different initializations have been considered. The final values of the optimized parameters are given in Table 3. The convergence of objective functions is shown in Fig. 3. It is clear that, unlike the conventional methods [12-16], the proposed approach finally leads to the optimal solution regardless the initial one. Therefore, the proposed approach can be used to improve the solution quality of classical methods.

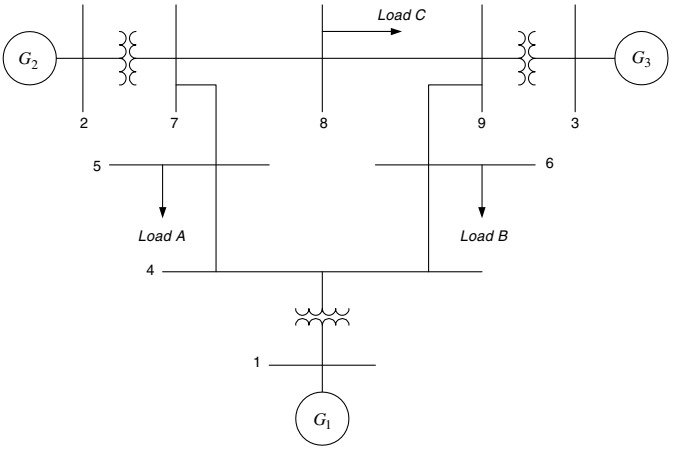


Fig. 2. Three-machine nine-bus power system

Table 1. Loads in pu on system 100-MVA base

Load	***Base Case***		***Case* 1**		***Case* 2**	
	P	*Q*	*P*	*Q*	*P*	*Q*
A	1.250	0.500	2.000	0.800	1.500	0.900
B	0.900	0.300	1.800	0.600	1.200	0.800
C	1.000	0.350	1.500	0.600	1.000	0.500

Table 2. Generator loadings in pu on the generator own base

Gen.	*Base Case*		*Case* 1		*Case* 2	
	P	*Q*	*P*	*Q*	*P*	*Q*
G_1	0.289	0.109	0.892	0.440	0.135	0.453
G_2	0.849	0.035	1.000	0.294	1.042	0.296
G_3	0.664	-0.085	1.000	0.280	1.172	0.298

Table 3. The optimal parameters of the proposed DEPSSs

	k	T_1	T_2	T_3	T_4
G_2	20.000	0.125	0.028	0.100	0.010
G_3	0.490	0.701	0.087	0.166	0.010

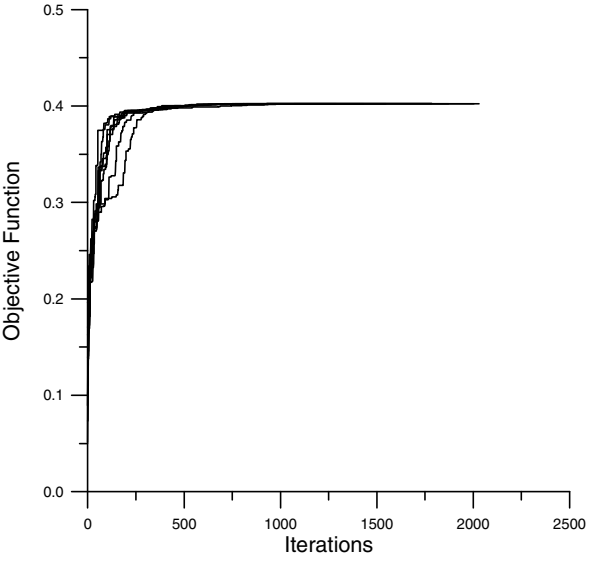


Fig. 3. Convergence of objective function with different initializations

4.2 Eigenvalue Analysis and Simulation Results

To assess the effectiveness and robustness of the proposed DEPSS over a wide range of loading conditions, two different cases designated as *case* 1 and *case* 2 are considered. The generator and system loading levels at these cases are given in Table 1 and Table 2 respectively. The electromechanical mode eigenvalues and the corresponding damping ratios without PSSs for all cases are given in Table 4. This table shows that the system has two local modes with frequencies of 1.44 Hz and 2.21 Hz in the base case. It is clear that these modes are poorly damped and some of them are unstable. The electromechanical mode eigenvalues and the

corresponding damping ratios with the proposed DEPSSs settings are given in Table 5. It is obvious that the electromechanical mode eigenvalues have been shifted to the left in *s*-plane and the system damping with the proposed DEPSSs is greatly improved and enhanced.

For further illustration, a 6-cycle three-phase fault disturbance at bus 7 at the end of line 5-7 is considered for the nonlinear time simulations. The speed deviations are shown in Fig. 4 with *base case*. The performance of the proposed DEPSS is compared to that of GA based PSS (GAPSS) given in [45]. It is clear that the proposed DEPSSs outperform the GAPSSs and provide good damping characteristics to low frequency oscillations and enhance greatly the dynamic stability of power systems.

Table 4. Eigenvalues and damping ratios without PSSs

Base Case	***Case* 1**	***Case* 2**
-0.01 ±*j* 9.07, *0.001*	-0.02±*j* 8.91, *0.002*	0.38 ±*j* 8.87, *-0.034*
-0.78 ±*j* 13.86, *0.056*	-0.52 ±*j* 13.83, *0.038*	-0.34 ±*j* 13.69, *0.025*

Table 5. Eigenvalues and damping ratios with the proposed DEPSSs

Base Case	***Case* 1**	***Case* 2**
-3.81 ±*j* 8.11, *0.427*	-2.98 ±*j* 6.73, *0.405*	-3.52 ±*j* 7.04, *0.447*
-6.41 ±*j* 13.70, *0.424*	-5.39 ±*j* 8.91, *0.517*	-4.04 ±*j* 9.17, *0.403*

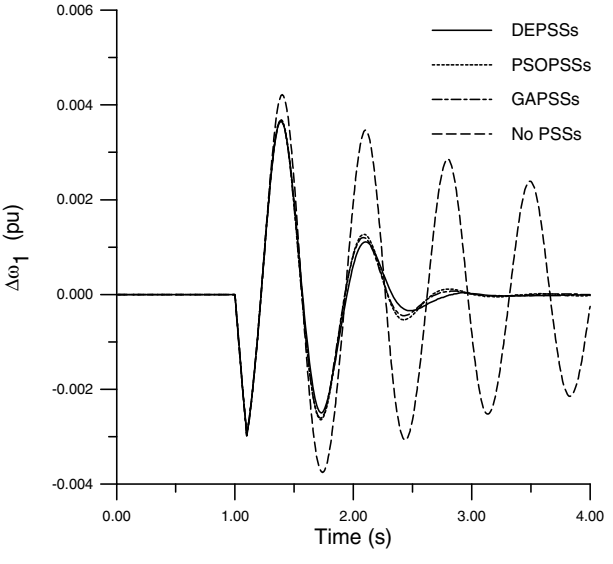


Fig. 4. System response to 6-cycle fault

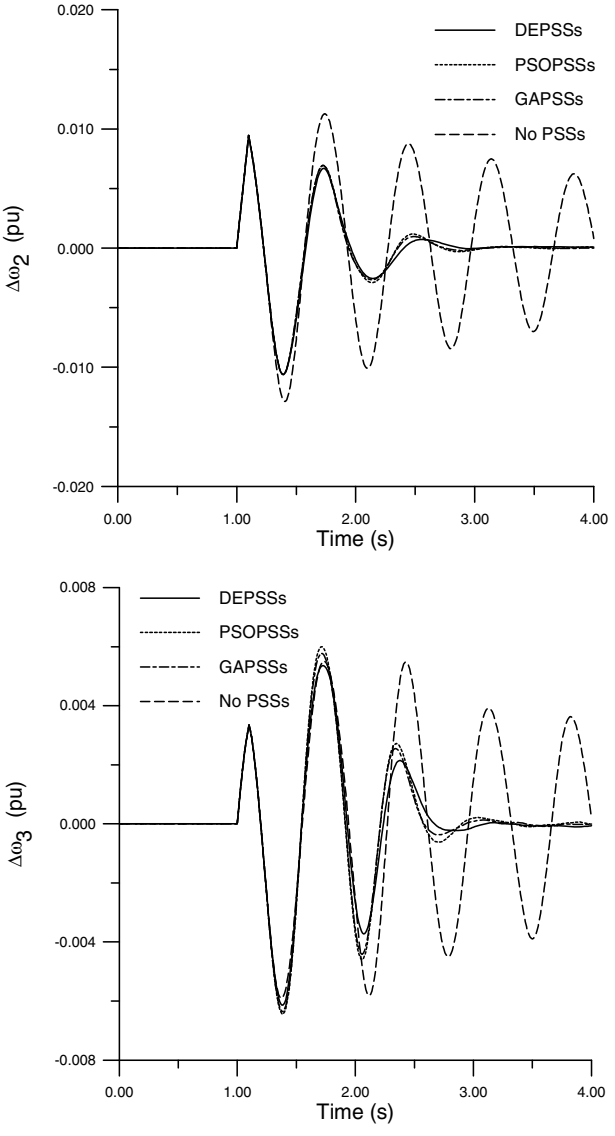


Fig. 4. (*Continued*)

5 Example 2: New England Power System

5.1 Test System and PSS Design

In this example, the 10-machine 39-bus New England power system shown in Fig. 5 is considered. Generator G_1 is an equivalent power source representing parts of the U.S.-Canadian interconnection system. Details of the system data are given in [46].

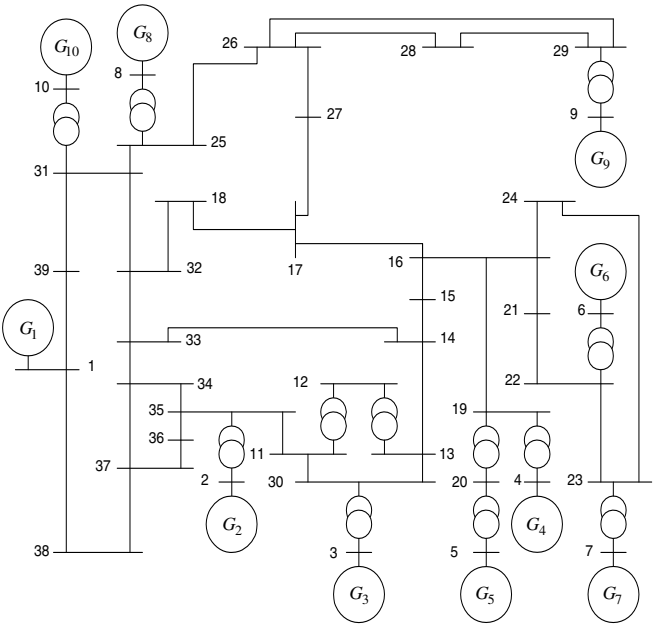


Fig. 5. Single line diagram for New England system

For illustration and comparison purposes, it is assumed that all generators except G_1 are equipped with PSSs. Hence, the optimized parameters are K_i, T_{1i}, and T_{3i}, i=2,3,…,10 *i.e.*, the number of optimized parameters is 27 in this example. DE algorithm has been applied to search for settings of these parameters so as to optimize the objective function considered. The final values of the optimized parameters are given in Table 6. The convergence of objective functions is shown in Fig. 6.

Table 6. The optimal parameters of the proposed DEPSSs

	k	T_1	T_2	T_3	T_4
G_2	48.430	0.336	0.045	0.546	0.027
G_3	22.915	0.915	0.038	0.556	0.052
G_4	47.778	0.682	0.019	0.397	0.042
G_5	48.795	0.132	0.030	0.241	0.036
G_6	49.346	0.157	0.059	0.690	0.063
G_7	7.254	0.186	0.058	0.194	0.035
G_8	13.329	0.662	0.035	1.272	0.025
G_9	21.098	0.314	0.049	0.159	0.079
G_{10}	17.019	1.054	0.024	1.049	0.015

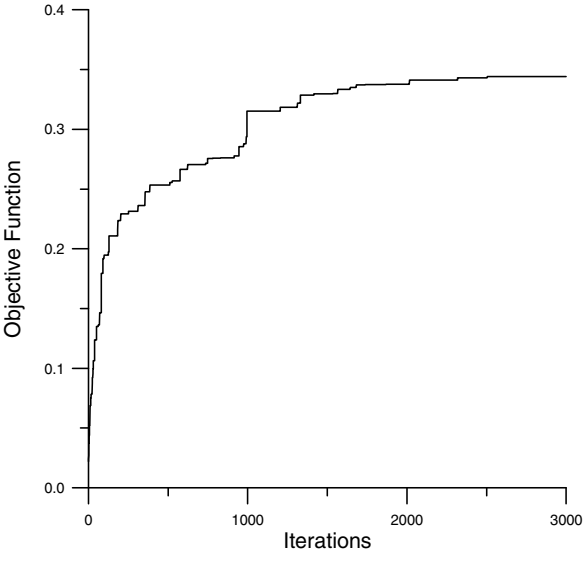


Fig. 6. Objective function convergence

5.2 Eigenvalue Analysis and Simulation Results

To demonstrate the effectiveness and robustness of the proposed DEPSSs under severe conditions and critical line outages, two different operating conditions in addition to the *base case* are considered. They can be described as: *Case 1*; outage of line 21-22; and *Case 2*; outage of line 14-15.

The electromechanical modes without PSSs for the three cases are given in Table 7. This table shows that the system has one interarea mode with frequency of o.64 Hz and eight local modes with frequencies ranging from 0.92 Hz to 1.54 Hz in the base case. It is clear that these modes are poorly damped and some of them are unstable. The electromechanical mode eigenvalues and the corresponding damping ratios with the proposed DEPSSs settings are given in Table 8. It can be seen that the electromechanical mode eigenvalues with the proposed DEPSSs have been shifted to the left in *s*-plane. It is obvious that the system damping is greatly improved and enhanced for all cases.

For time-domain simulations, a 6-cycle three-phase fault at bus 29 at the end of line 26-29 has been applied to demonstrate the effectiveness and the robustness of the proposed DEPSSs.

The performance of the proposed DEPSSs is compared to that of GAPSSs given in [19] and PSO-based PSSs given in [22]. Due to space limitations, only the speed deviations of G_8 and G_{10} are shown in Fig. 7. It is clear that the system performance with the proposed DEPSSs is much better than that of GAPSSs and PSOPSSs and the oscillations are damped out much faster. In addition, the proposed DEPSSs are quite efficient to damp out the local modes as well as the interarea modes of oscillations. This illustrates the potential and superiority of the proposed design approach to get optimal set of PSS parameters.

Table 7. Eigenvalues and damping ratios without PSSs

Base Case	***Case* 1**	***Case* 2**
0.191 ±*j* 5.808, *-0.033*	**0.195 ±*j* 5.716, *-0.034***	**0.152 ±*j* 5.763, *-0.026***
0.088 ±*j* 4.002, *-0.022*	**0.121 ±*j* 3.798, *-0.032***	**0.095 ±*j* 3.837, *-0.025***
-0.028 ±*j* 9.649, *0.003*	**0.097 ±*j* 6.006, *-0.016***	**0.033 ±*j* 6.852, *-0.005***
-0.034 ±*j* 6.415, *0.005*	-0.032 ±*j* 9.694, *0.003*	-0.026 ±*j* 9.659, *0.003*
-0.056 ±*j* 7.135, *0.008*	-0.104 ±*j* 8.015, *0.013*	-0.094 ±*j* 8.120, *0.012*
-0.093 ±*j* 8.117, *0.011*	-0.109 ±*j* 6.515, *0.017*	-0.100 ±*j* 6.038, *0.017*
-0.172 ±*j* 9.692, *0.018*	-0.168 ±*j* 9.715, *0.017*	-0.171 ±*j* 9.696, *0.018*
-0.220 ±*j* 8.013, *0.027*	-0.204 ±*j* 8.058, *0.025*	-0.219 ±*j* 8.000, *0.027*
-0.270 ±*j* 9.341, *0.029*	-0.250 ±*j* 9.268, *0.027*	-0.259 ±*j* 9.320, *0.028*

Table 8. Eigenvalues and damping ratios with the Proposed DEPSSs

Base Case	***Case* 1**	***Case* 2**
-1.279±*j* 7.929, *0.159*	-1.307±*j* 6.666, *0.192*	-1.301 ±*j* 9.174, *0.140*
-1.326 ±*j* 8.467, *0.155*	-1.321 ±*j* 9.170, *0.143*	-1.309 ±*j* 3.719, *0.332*
-1.359 ±*j* 9.165, *0.147*	-1.495 ±*j* 8.282, *0.178*	-1.370 ±*j* 6.551, *0.205*
-1.667 ±*j* 9.877, *0.166*	-1.589 ±*j* 3.628, *0.401*	-1.486 ±*j* 8.271, *0.177*
-1.713 ±*j* 3.474, *0.442*	-1.699 ±*j* 10.007, *0.167*	-1.668 ±*j* 10.013, *0.164*
-1.791 ±*j* 10.279, *0.172*	-1.733 ±*j* 10.287, *0.166*	-1.699 ±*j* 10.291, *0.163*
-1.885 ±*j* 9.626, *0.192*	-1.763 ±*j* 4.108, *0.394*	-1.857 ±*j* 3.895, *0.430*
-2.337 ±*j* 4.451, *0.465*	-1.862 ±*j* 9.754, *0.188*	-1.872 ±*j* 9.737, *0.189*
-2.740 ±*j* 10.579, *0.251*	-2.601 ±*j* 10.333, *0.244*	-2.538 ±*j* 10.292, *0.239*

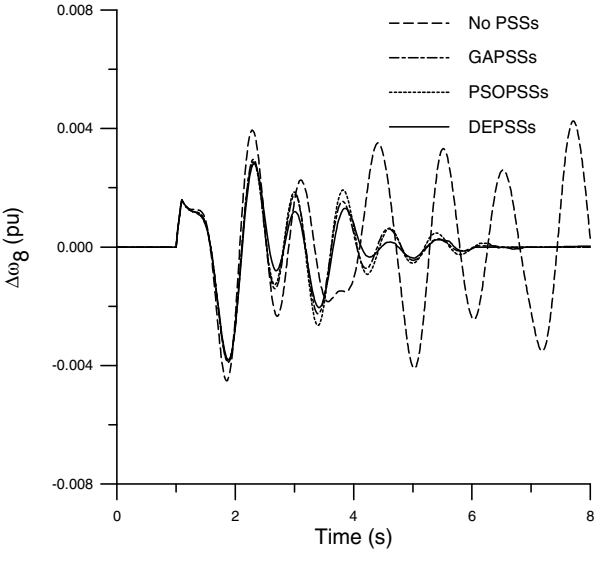


Fig. 7. System response with 6-cycle 3-phase fault disturbance

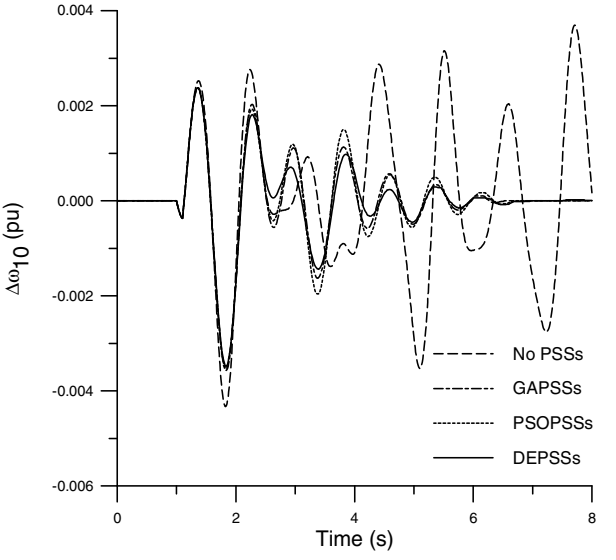


Fig. 7. (*Continued*)

For completeness and clear perceptiveness about the system response, two performance indices that reflect the settling time and overshoots are introduced and evaluated. These indices are defined as

$$PI_1 = \sum_{i=1}^{n} \int_{t=0}^{t=t_{sim}} (t\Delta\omega_i)^2 \, dt \tag{13}$$

$$PI_2 = \sum_{i=1}^{n} \int_{t=0}^{t=t_{sim}} (\Delta\omega_i)^2 \, dt \tag{14}$$

where n is the number of machines and t_{sim} is the simulation time. The values of these indices with the disturbances considered are given in Table 9. It is clear that the values of these indices with the proposed DEPSSs are much smaller compared to both PSOPSSs and GAPSSs. This demonstrates that the settling time and the speed deviations of all units are much reduced by applying the proposed DEPSSs. Additionally, the proposed DEPSSs provide good damping characteristics and outperform PSOPSSs and GAPSSs.

Table 9. Performance indices with different systems

	PI_1				PI_2			
	No PSSs	**GA-PSSs**	**PSO-PSSs**	**DE-PSSs**	**No PSSs**	**GA-PSSs**	**PSO-PSSs**	**DE-PSSs**
3-machine	3.669	0.207	0.224	0.191	1.739	0.580	0.596	0.558
10-machine	194.410	1.451	1.696	1.581	10.402	1.011	1.062	1.056

Acknowledgments. The author acknowledges the support of King Fahd University of Petroleum & Minerals, Saudi Arabia, via funded project # IN090019.

Appendix

In this work, the *i*-th machine model is given as follows.

$$\dot{\delta}_i = \omega_b(\omega_i - 1) \tag{A1}$$

$$\dot{\omega}_i = (T_{mi} - T_{ei} - D_i(\omega_i - 1))/M_i \tag{A2}$$

$$\dot{E}'_{q\,i} = (E_{fdi} - (x_{di} - x'_{di})i_{di} - E'_{qi})/T'_{doi} \tag{A3}$$

$$\dot{E}'_{d\,i} = (-(x_{qi} - x'_{qi})i_{qi} - E'_{di})/T'_{qoi} \tag{A4}$$

$$\dot{E}_{fd\,i} = (K_{ai}(V_{refi} - V_i - U_i) - E_{fdi})/T_{ai} \tag{A5}$$

$$T_{ei} = E'_{di}i_{di} + E'_{qi}i_{qi} - (x'_{qi} - x'_{di})i_{di}i_{qi} \tag{A6}$$

where *d* and *q* refer to the direct and quadrature axes respectively. δ is the rotor angle; ω is rotor speed; E'_q and E'_d are the internal voltages behind x'_d and x'_q respectively; E_{fd} is the equivalent excitation voltage; T_e is the electric torque; T'_{do} and T'_{qo} are time constants of excitation circuit; K_a is the regulator gain; T_a is the regulator time constant.

References

[1] Anderson, P.M., Fouad, A.A.: Power System Control and Stability. Iowa State Univ. Press, Ames (1977)

[2] Sauer, P.W., Pai, M.A.: Power system Dynamics and Stability. Prentice-Hall, Englewood Cliffs (1998)

[3] de Mello, F.P., Concordia, C.: Concepts of Synchronous Machine Stability as Affected by Excitation Control. IEEE Trans. PAS 88, 316–329 (1969)

[4] Xia, D., Heydt, G.T.: Self-Tuning Controller for Generator Excitation Control. IEEE Trans. PAS 102, 1877–1885 (1983)

[5] Cao, Y., Jiang, L., Cheng, S., Chen, D., Malik, O.P., Hope, G.S.: A Nonlinear Variable Structure Stabilizer for power System Stability. IEEE Trans. on EC 9(3), 489–495 (1994)

[6] Hiyama, T., Sameshima, T.: Fuzzy Logic Control Scheme for On-Line Stabilization of multimachine Power System. Fuzzy Sets and Systems 39, 181–194 (1991)

[7] Abido, M.A., Abdel-Magid, Y.L.: A hybrid Neuro-Fuzzy Power System Stabilizer for Multimachine Power Systems. IEEE Trans. on PWRS 13(4), 1323–1330 (1998)

[8] Abido, M.A., Abdel-Magid, Y.L.: Hybridizing Rule-Based Power System Stabilizers with Genetic Algorithms. IEEE Trans. on PWRS 14(2), 600–607 (1999)
[9] Larsen, E., Swann, D.: Applying Power System Stabilizers. IEEE Trans. PAS 100(6), 3017–3046 (1981)
[10] Tse, G.T., Tso, S.K.: Refinement of Conventional PSS Design in Multimachine System by Modal Analysis. IEEE Trans. PWRS 8(2), 598–605 (1993)
[11] Kundur, P., Klein, M., Rogers, G.J., Zywno, M.S.: Application of Power System Stabilizers for Enhancement of Overall System Stability. IEEE Trans. PWRS 4(2), 614–626 (1989)
[12] Abe, S., Doi, A.: A New Power System Stabilizer Synthesis in Multimachine Power Systems. IEEE Trans. PAS 102(12), 3910–3918 (1983)
[13] Arredondo, J.M.: Results of a Study on Location and Tuning of Power System Stabilizers. Int. J. Electrical Power & Energy Systems 19(8), 563–567 (1997)
[14] Lim, C.M., Elangovan, S.: Design of Stabilizers in Multimachine Power Systems. IEE Proc. 132(3), Pt. C, 146–153 (1985)
[15] Chen, C.L., Hsu, Y.Y.: Coordinated Synthesis of Multimachine Power System Stabilizer using an Efficient Decentralized Modal Control Algorithm. IEEE Trans. PWRS 2(3), 543–551 (1987)
[16] Yu, Y.N., Li, Q.: Pole-Placement Power System Stabilizers Design of an Unstable Nine-Machine System. IEEE Trans. PWRS 5(2), 353–358 (1990)
[17] Maslennikov, V.A., Ustinov, S.M.: The Optimization Method for Coordinated Tuning of Power System Regulators. In: Proc. 12th Power System Computation Conference PSCC, Dresden, pp. 70–75 (1996)
[18] Taranto, G.N., Falcao, D.M.: Robust Decentralised Control Design Using Genetic Algorithms in Power System Damping Control. IEE Proc. Genet. Transm. Distrib. 145(1), 1–6 (1998)
[19] Abdel-Magid, Y.L., Abido, M.A., Al-Baiyat, S., Mantawy, A.H.: Simultaneous Stabilization of Multimachine Power Systems Via Genetic Algorithms. IEEE Trans. PWRS 14(4), 1428–1439 (1999)
[20] Abido, M.A.: A novel approach to conventional power system stabilizer design using tabu search. International Journal of Electrical Power & Energy Systems 21(6), 443–454 (1999)
[21] Abido, M.A.: Robust Design of Multimachine Power System Stabilizers Using Simulated Annealing. IEEE Trans. on Energy Conversion 15(3), 297–304 (2000)
[22] Abido, M.A.: Optimal Design of Power System Stabilizers Using Particle Swarm Optimization. IEEE Trans. on Energy Conversion 17(3), 406–413 (2002)
[23] Fogel, D.B.: Evolutionary Computation Toward a New Philosophy of Machine Intelligence. IEEE Press, Los Alamitos (1995)
[24] Differential Evolution web site, `http://www.icsi.erkeley.edu/~storn/code.html`
[25] Storn, R., Price, K.: Differential evolution – a simple and efficient adaptive scheme for global optimization over continuous spaces. Technical Report TR-95-012, `ftp://ftp.icsi.berkeley.edu/pub/techreports/1995/tr-95-012.pdf`
[26] Storn, R., Price, K.: Minimizing the real functions of the ICEC 1996 contest by differential evolution. In: Proceedings of IEEE International Conference on Evolutionary Computation, May 20-22, pp. 842–844 (1996)

[27] Storn, R.: On the usage of differential evolution for function optimization. In: 1996 Biennial Conference of the North American Fuzzy Information Processing Society, NAFIPS, June 19-22, pp. 519–523 (1996)
[28] Das, S., Abraham, A., Konar, A.: Particle Swarm Optimization and Differential Evolution Algorithms: Technical Analysis, Applications and Hybridization Perspectives, http://www.softcomputing.net/aciis.pdf
[29] Karaboga, D., Ökdem, S.: A Simple and Global Optimization Algorithm for Engineering Problems: Differential Evolution Algorithm. Turkish Journal Electrical Engineering 12(1) (2004)
[30] Storn, R., Price, K.: Differential Evolution, A Simple and Efficient Heuristic Strategy for Global Optimization over Continuous Spaces. Journal of Global Optimization 11, 341–359 (1997)
[31] Vesterstrom, J., Thomsen, R.: A comparative study of differential evolution, particle swarm optimization and evolutionary algorithms on numerical benchmark problems. In: Congress on Evolutionary Computation, CEC 2004, June 19-23, vol. 2, pp. 1980–1987 (2004)
[32] Yang, Z., Tang, K., Yao, X.: Differential Evolution for High-Dimensional Function Optimization. In: IEEE Congress on Evolutionary Computation (CEC 2007), pp. 3523–3530 (2007)
[33] Lampinen, J.: A Bibliography of Differential Evolution Algorithm, http://www.lut.fi/~jlampine/debiblio.htm
[34] Su, C.T., Lee, C.S.: Network Reconfiguration of Distribution Systems using Improved Mixed-Integer Hybrid Differential Evolution. IEEE Transactions on Power Delivery 18(3), 1022–1027 (2003)
[35] Chiou, J.P., Chang, C.F., Su, C.T.: Ant Direction Hybrid Differential Evolution for Solving Large Capacitor Placement Problems. IEEE Transactions on Power System 19(4), 1794–1800 (2004)
[36] Mayer, D.G., Kingborn, B.P., Archer, A.A.: Differential Evolution – An Easy and Efficient Evolutionary Algorithm for Model Optimization. Agric. Syst. 83, 315–328 (2005)
[37] Kapadi, M.D., Gudi, R.D.: Optimal Control of Fed-Batch Fermentation Involving Multiple Feeds using Differential Evolution. Process Biochemistry 39(11), 709–1721 (2004)
[38] Babu, B.V., Chakole, P.G., Mubeen, J.H.S.: Differential Evolution Strategy for Optimal Design of Gas Transmission Network, http://www.vsppub.com
[39] Balamurugan, R., Subramanian, S.: Self-Adaptive Differential Evolution Based Power Economic Dispatch of Generators with Valve-Point Effects and Multiple Fuel Options. International Journal of Computer Science and Engineering 1(1), 10–17
[40] Liang, C.H., Chung, C.Y., Wong, K.P., Duan, X.Z., Tse, C.T.: Study of Differential Evolution for Optimal Reactive Power Flow. IET Gener. Transm. Distrib. 1(2), 253–260 (2007)
[41] Thomas, P., Vernon, D.: Image Registration by Differential Evolution. In: Proceedings of the Irish Machine Vision and Image Processing Conference, pp. 221–225 (1997)
[42] Storn, R.: Differential Evolution Design of an IIR-filter with Requirements of Magnitude and Group Delay. In: Proceedings of the IEEE Conference on Evolutionary Computation, pp. 268–273 (1996)
[43] Storn, R.: System Design by Constraint Adaptation and Differential Evolution. IEEE Transactions on Evolutionary Computation 3(1), 22–34 (1999)

[44] Daniela, Z.: A Comparative Analysis of Crossover Variants in Differential Evolution". In: Proceedings of the International Multiconference on Computer Science and Information Technology, pp. 171–181 (2007)
[45] Abido, M.A.: Parameter optimization of multimachine power system stabilizer using genetic local search. International Journal of Electrical Power & Energy Systems 23(8), 785–794 (2001)
[46] Pai, M.A.: Energy Function Analysis for Power System Stability. Kluwer Academic Publishers, Dordrecht (1989)

An AIS-ACO Hybrid Approach for Multi-Objective Distribution System Reconfiguration

A. Ahuja, S. Das, and A. Pahwa

Abstract. This work proposes a hybrid algorithm based on artificial immune systems and ant colony optimization for distribution system reconfiguration, which is formulated as a multi-objective optimization problem. The algorithm maintains a population of candidate solutions called antibodies. The search space is explored by means of the hypermutation operator that perturbs existing antibodies to produce new ones. A table of pheromones is used to reinforce better edges during hypermutation. An added innovation is the use of the pheromones to obtain quick solutions to restore the distribution system under contingency situations. The hybrid approach has been successfully implemented on two test networks. The results obtained demonstrate the efficacy of the algorithm.

1 Introduction

Distribution Systems deliver electric power from transmission system to customers. They undergo different loading patterns at different times during a typical 24-hour period. In addition, the loading patterns change by the day of the week and by the season. Load on the feeders of a distribution system is generally a combination of industrial, commercial, residential and lighting load. Substation transformers and feeders undergo peak loading at different times of the day, and therefore the distribution system becomes heavily loaded at certain times in a day, and lightly loaded at others. This variation of load is detrimental to the operating conditions of the network and leads to high real losses, poor voltage profile, unbalanced loading on the substation transformers and feeders, and reduction in reserve capacity of transformers and feeders. High real losses on the distribution network mean revenue loss for the utility; poor voltage profile means poor power quality at the customer's premises, which in turn poses a threat to the reliability of the system and credibility of the utility. Low reserve capacity and unbalanced loading on transformers and feeders drastically affect the systems capability to restore power to disrupted customers under contingencies like faults, which effectively means reliability of the distribution system is compromised. Thus, formulation of a strategy that can simultaneous optimize multiple objectives like minimization of real losses, improvement of voltage profile and balancing of load on transformers and feeders holds a lot of promise and is of prime importance for a utility to maximize its revenue and improve its credibility in today's competitive market.

Distribution Systems are structurally meshed but are operated in a radial configuration. They use a set of switches, called sectionalizing-switches, that remain normally closed, and another set of switches, called tie-switches, that remain normally open. Changing the status of these switches can change the topology of a distribution network. The process of changing the topology of distribution systems by

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 19–73.
springerlink.com

altering the open/closed status of the switches is called Distribution System Reconfiguration and can be used to improve the operating conditions of the system. (Note: In this work, the term Distribution Network Reconfiguration has also been used in place of Distribution System Reconfiguration. Both mean the same thing.)

With the advancement in automation of distribution network, it has become possible to automate all the switches in a distribution network such that their open/closed status can be changed quickly from a remote location without any human intervention. This makes possible a near real time change in the network topology, which brings about a fast transfer of load from heavily loaded feeders to lightly loaded feeders, resulting in improvement of the operational state of the network. Thus the problem of network reconfiguration can be defined as the search of a feasible network topology that satisfies all the operational constraints like capacity constraint, voltage constraint and simultaneously optimizes multiple objectives like minimization of real losses on the network and balancing the load on the transformers and feeders. In addition to this, network reconfiguration can also be used during contingencies for restoring power to as many customers as possible and also for system expansion and planning purposes. Thus it is extremely promising to devise an effective solution methodology that could search for a network topology that simultaneously optimizes the objectives mentioned above and hence taps into the benefits offered by distribution network reconfiguration. This has been the prime motivation behind proposing a multi-objective optimization framework for distribution network reconfiguration problem.

Mathematically, distribution system reconfiguration problem is a complex, combinatorial, constrained optimization problem. The complexity of the problem arises from the fact that distribution network topology has to be radial and power flow constraints are non-linear in nature. A distribution system can be modeled as an undirected graph with a set of edges and nodes, where nodes correspond to the buses and edges correspond to the lines between the buses. The search for the 'optimal' network topology can then be formulated as a search for quadratic minimum spanning tree problem (q-MST). And it has been shown in [1] that quadratic minimum spanning tree (q-MST) problem is NP-hard [2], which means it is very difficult to find the optimal solution in polynomial time. Thus it is extremely hard for exact algorithms such as branch and bound, branch and cut, which assume the constraints to be linear, to solve a problem of this complexity. And algorithms based on problem specific heuristics are greedy in nature and explore only a very narrow region of the search space and thus have a tendency of getting stuck into locally optimal solutions. Above all, it is extremely tough to solve the reconfiguration problem for simultaneous optimization of multiple objectives using exact algorithms or heuristics-based algorithms. Recently, algorithms inspired from nature like Genetic Algorithm [3], Ant Colony Optimization [4], Particle Swarm Optimization [5] have been successfully applied for simultaneous optimization of multiple objectives for various NP-hard combinatorial problems like Traveling Salesman Problem (TSP), Quadratic Assignment Problem (QAP) with linear/non linear constraints. Also, these algorithms are population-based approaches. So instead of just one solution, a set of optimal solutions, based on Pareto Optimality [6], is obtained. These observations have

inspired the use of nature-based algorithms for solving multi-objective Distribution System Reconfiguration problem in this work.

Till now researchers have predominantly modeled the reconfiguration problem as a single objective optimization problem only; primarily minimizing real losses on the feeders subject to operational constraints like voltage deviation, line and transformer capacity constraints etc [7-12]. This work proposes a multi-objective framework for distribution network reconfiguration problem, based on Pareto Optimality, where objectives like minimization of real losses and transformer load balancing are simultaneously optimized. Moreover, the philosophy behind the optimization methods used so far in literature has been to search for a solution in a narrow region of the search space only, which means finding network topologies that do not violate any operational constraints. Formulation of operational constraints such as voltage deviation as objectives in the multi-objective framework, is such that the search for a solution is not just limited to a narrow region of the search space, but is extended to other regions as well. The inspiration behind this approach is the fact that there could be certain solutions that optimize the primary objective function (like loss minimization) better than good feasible solutions but violate the constraints just by a small margin. Solutions like these can be useful for distribution system planning and expansion purposes. Hence search for good 'infeasible' solutions seems useful. The concept of Pareto Optimality has been used to qualify the solutions in multi-objective optimization framework.

The second contribution of this work is the development of nature-inspired algorithm to effectively solve this NP hard, discrete optimization problem. A novel hybrid algorithm has been proposed to solve multi-objective reconfiguration problem. This algorithm combines concepts from Artificial Immune Systems (AIS) and Ant Colony Optimization (ACO). Artificial Immune System uses random mutation as its search mechanism, which makes the search process too exploratory and thus the convergence to a good solution is slow. To improve the convergence characteristics, a pheromone-based mutation has been used in AIS algorithm.

Furthermore, distribution network reconfiguration problem has been expanded to incorporate issues related to restoration of power during contingencies. Contingency in a distribution system could be an occurrence of a fault on a feeder, on a transformer at the substation or occurrence of a sustained interruption due to failure of any component of the system. The severity of a fault primarily depends on the number of customers affected by the fault and the minimum amount of time necessary to restore the faulted part of the system. Thus the primary objective of energy restoration, after a fault has been cleared, is to supply power to as many disrupted customers as possible and as fast as possible. This is achieved by changing the topology of the network such that faulted part of the network can be fed power from unfaulted part of the network. So, mathematically energy restoration problem is equivalent to a constrained distribution reconfiguration problem, where the additional constraint, apart from the operational constraints, is that the faulted component of the network will always be open till the time it is repaired and restored. The new hybrid AIS-ACO framework makes it possible to use the information stored in pheromones, which they gather while searching for a solution for loss minimization, for solving energy restoration problem during contingencies.

Finally, some test cases have been presented to illustrate and validate the effectiveness of the algorithm.

The earliest work reported in literature on distribution network reconfiguration dates back to 1975 proposed by Merlin and Back [7]. The technique proposed by Merlin and Back exploited the radial topology of a distribution network and was used for minimizing the real losses on the network. Since then researchers have proposed several categories of network reconfiguration techniques for power loss minimization.

The simplest and a brute-force method would be to search through all the feasible configurations of the system and choose whichever optimizes the objective function to the maximum while staying within the operational constraints. Since a typical distribution system may have hundreds of switches, a combinatorial analysis of all possible options, as done in [13], is a not a practical proposition. The radiality constraint and the discrete nature of the switch values prevent the use of classical optimization techniques to solve the reconfiguration problem. Therefore, most of the algorithms in literature are based on heuristic search techniques. Civanlar *et al.* [14] suggested a branch-exchange type algorithm, where a simple formula has been derived to determine how a branch exchange affects the losses. In [8], Baran *et al.* have modeled reconfiguration problem for loss minimization and load balancing as an integer-programming problem. In a variation of switch exchange approach [9], a piecewise quadratic problem is solved to find the optimum tie switch in a loop. In Shirmohammadi and Hong [10], the solution method starts with a meshed distribution system obtained by considering all switches closed. Then, the switches are opened successively to eliminate the loops. Goswami and Basu [11] report a heuristic algorithm that is based on the concept of optimum flow pattern that is determined by using a power flow program. The optimum flow pattern of a single loop formed by closing a normally open switch is found out, and this flow pattern is established in the radial network by opening a closed switch. This procedure is repeated until the minimum loss configuration is obtained. McDermott *et al.* [12] proposed a heuristic constructive algorithm that starts with all maneuverable/tie switches open, and at each step, the switch that results in the least increase in the objective function is closed. The objective function is defined as the incremental losses divided by incremental load served. Recently Gomes *et al.* [15] have proposed a heuristic based algorithm that starts with the system in a meshed status with all switches closed. Switches are opened one by one based on the calculation of the minimum total system loss using a load flow program. In Ramos *et al.* [16], a new path-to-node based modeling for distribution system has been proposed. They have used this model with Mixed Integer Programming and Genetic algorithm for solving reconfiguration problem for loss minimization. Schmidt *et al.* [17] have proposed a method for loss minimization based on standard Newton method.

The other class of approaches for solving distribution system reconfiguration problem for loss minimization has been based on techniques from Artificial Intelligence. Momoh *et al.* [18] have proposed a method based on artificial neural network for loss reduction and voltage deviation minimization. In Kim *et al.* [19],

Kagan *et al.* [20], the authors have employed methods based on Genetic Algorithms. Jeon et al. [21] have used simulated annealing and tabu search for solving network reconfiguration for loss minimization. Recently, Delbem et al. [22] have combined concepts from graph theory and evolutionary algorithms for solving reconfiguration problem. In Chiou *et al.* [23] an algorithm based on differential evolution has been proposed for solving reconfiguration. In Ahuja and Pahwa [24], ant colony optimization (ACO) technique has been proposed for loss minimization. In Jin *et al.* [25], a method based on particle swarm optimization has been proposed.

All of the above mentioned works have modeled the reconfiguration problem as a single objective optimization problem, primarily minimization of real losses. Only some researchers [26-27] have formulated the reconfiguration problem as a multi-objective optimization problem. But none of them have used the idea of Pareto Optimality in their multi-objective framework. In [26] Roytelman *et al.* have proposed simultaneously optimizing multiple objectives through reconfiguration. But they combined all the objectives into one single objective by taking a weighted sum of different objectives. In [27], Hsiao has proposed a similar multi objective approach for feeder reconfiguration based on Evolution Programming. He uses fuzzy membership functions for all the objectives considered and searches for a solution that matches the predefined fuzzy membership values. This approach, in effect, is equivalent to the approach proposed by Roytelman *et al* [26]. The only difference is that objective values have been replaced by fuzzy membership values; but the search process is guided by predefined fuzzy membership value of the objectives in the same manner as weights guide the search process in [26]. The problem with these approaches is that the search process explores only a narrow region of the search space, which is dependent on the set of weights used (or the set of predefined fuzzy membership values). Also, only a single solution is produced during the search process. This could limit the choice for the final solution, if for instance, inappropriate set of weights or fuzzy membership functions were used, which might result in undesired or sub optimal solution. And if more than one solution is desired, the search process has to be performed as many times as the number of solutions desired, which could be computationally infeasible for large real systems. The details of the shortcomings of this approach have been elaborated in [28]. These observations motivated the use of the concept of Pareto Optimality for multi-objective framework for distribution network reconfiguration, in which, instead of just one optimal solution, the set of all optimal solutions is preserved.

Section 2 presents the formal definition of Pareto Optimality and explains the multi-objective framework used. Further it presents the mathematical formulation for multi-objective distribution system reconfiguration problem. Section 3 describes the theoretical basis of Artificial Immune System and Ant Colony Optimization and presents the AIS-ACO hybrid algorithm. Section 4 explains the solution methodology used for solving multi-objective reconfiguration problem. Section 5 presents the results obtained on the test cases.

2 Problem Formulation

This section describes the multi-objective framework used for distribution network reconfiguration problem. It also gives a detailed explanation of the mathematical formulation used for the multi-objective reconfiguration problem and the graphical model used to represent a distribution system.

2.1 Distribution System Modeling

The most intuitive way of representing a distribution network is by a graph. A lot of researchers have adopted this method of representation, for example [29]. In general, a distribution network is represented by an undirected graph $G = \{V, E\}$ with a nonempty set of nodes V representing buses and a set of edges E representing the lines between the buses. As mentioned in section 1.1, distribution systems are structurally meshed but are operated in a radial configuration. Radial operation basically means that for every bus in the system there is just one unique path from substation to that bus and there are no loops in the system. This in effect means that every bus is fed power from just one substation. On a graph, the radial configuration of a distribution system is equivalent to a spanning tree. A spanning tree of a graph is defined as a subgraph that connects all the nodes of the graph to the root node such that there are no cycles in the graph. Consider the small three-feeder distribution system [23] given ahead. As defined in the earlier, distribution network reconfiguration problem is a search for a network topology that simultaneously optimizes multiple objectives like minimization of real losses on the system, balancing of loads on the substation transformers etc., while satisfying operational constraints like staying within voltage deviation limit and staying within line and transformer capacity limit. Thus, in the context of a graphical model of a

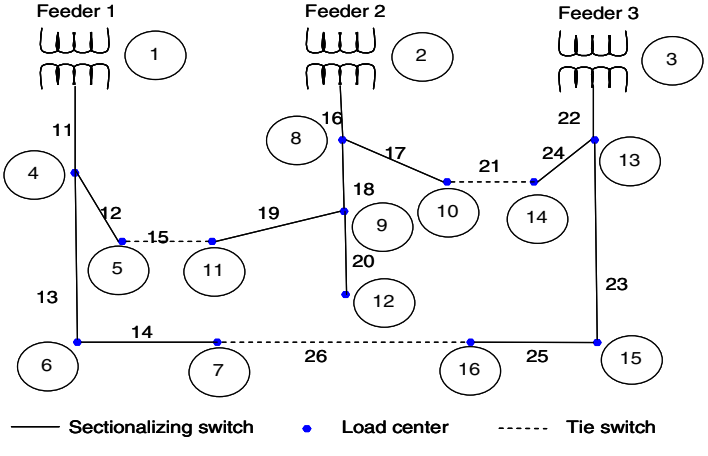


Fig. 1. A small three-feeder distribution system [23]

distribution network, reconfiguration problem is the search for a spanning tree of the graph that achieves simultaneous optimization of these objectives. This entails use of a multi objective framework for solving distribution network reconfiguration problem. The next subsection describes the multi-objective optimization framework in detail based on the work of Coello Coello [6].

The graphical model for this small system is obtained by replacing substation buses with root nodes. The Figure 2. below shows the graphical model for this system.

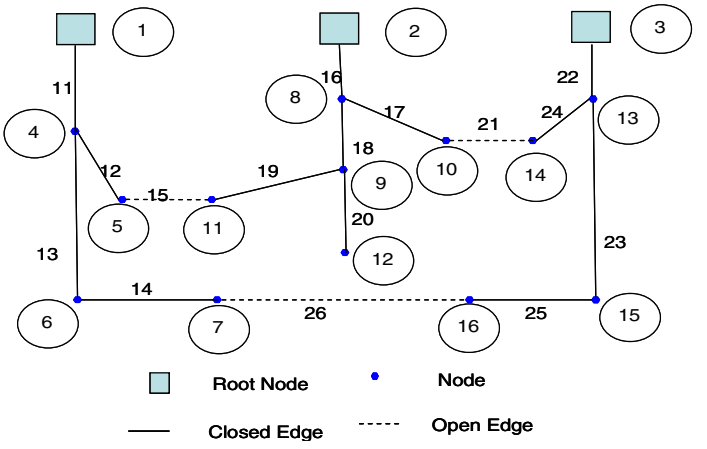


Fig. 2. Graphical model of the three-feeder system

2.2 Multi-objective Optimization Framework

Multi-objective optimization is defined as the problem of finding a vector of decision variables that satisfies all the constraints and simultaneously optimizes a vector function whose elements represent the objective functions. These objective functions are a mathematical description of the performance measure for a problem and are usually in conflict with each other.

Mathematically, it can be stated as finding a vector $x^* = [x_1^*, x_2^*, ...x_n^*]^T$ that will satisfy the m inequality constraints $g_i(x) \geq 0$, where $i = 1, 2, ..., m$, the p equality constraints $h_j(x) = 0$ where $j = 1, 2, ..., p$ and simultaneously optimizes the vector function $f(x) = [f_1(x), f_2(x), ...f_k(x)]^T$ where $x = [x_1, x_2, ...x_n]^T$ is the vector of decision variables. It is not often the case that there is a single point that optimizes all the objective functions simultaneously for a multi-objective optimization problem. Therefore, the search process is normally directed towards finding trade-offs among the objective functions, rather than a single globally optimal solution. The concept of optimality for multi-objective problems is therefore different. The most prevalent notion of optimality was originally proposed

by Edgeworth [30], and later generalized by Pareto [31]. Some authors call it *Edgeworth-Pareto* optimum, but the commonly accepted term for it is *Pareto Optimum.*

Without loss of generality, if we assume that the optimization problem involves minimization of the objectives functions, then a vector of decision variables $x^* \in \Omega$, where Ω is the set of all x that satisfy all constraints, is said to be Pareto Optimal if and only if there does not exist another $x \in \Omega$ such that $f_i(x) \leq f_i(x^*)$ for all $i = 1,.....,k$ and $f_i(x) < f_i(x^*)$ for at least one i. This means that there is no other decision vector that has any of the objective values better than objective values of vector x.

When comparing two decision vectors, a vector u is said to dominate vector v if and only if all objective values of u are less than or equal to the objective values of v and there exists at least one objective whose value is less for u than v. This basically means that u dominates v when it is better than v in at least one objective. The set of solution vectors that are not dominated by any other vectors in the decision space is called non-dominated solutions or Pareto solutions. Any optimization algorithm aims to find the Pareto solution set for a problem. When plotted in the objective space, the surface that the Pareto solutions generate is called Pareto Front.

The Pareto Optimality framework described in this section has been used for solving multi-objective distribution system reconfiguration problem. The next section presents the multi-objective distribution system reconfiguration problem in detail.

2.3 Multi-objective Reconfiguration Problem

Distribution Network Reconfiguration can be used to simultaneously optimize several objectives like minimization of real loss, load balancing on transformers and feeders. Moreover, the information gained from optimizing these objectives can be used for restoration during contingencies and can also be used for distribution planning and expansion purposes. Specifically, the methodology provides a feasible solution very quickly for restoration of power to the disrupted sections following a fault and isolation of the faulted section. Quick restoration decreases the duration of outages to customers and thus increases reliability.

In this work, distribution network reconfiguration problem is solved for simultaneous minimization of real loss, balancing of load on transformers and minimization of voltage drop. Mathematically, the multi-objective distribution system reconfiguration problem can be defined as:

$Minimize \quad F(G) = [f_1(G), f_2(G), f_3(G)]^T$ where $f_1(G)$ represent real losses, $f_2(G)$ represents unbalance in transformer loading and $f_3(G)$ represents deviation of voltage magnitude from 1 per unit. These functions are described in detail below.

1) Minimization of Real Loss: For a given configuration G of the distribution network, Total Real Loss is defined as:

$f_1(G) = \sum_i I_i^2 * r_i$, where $i = \{1, 2, ...NumberOfConnectedBranches\}$ and

I_i, is current (in per unit) in the i^{th} connected branch;

r_i, is resistance (in per unit) of the i^{th} connected branch.

2) Transformer Load Balancing: Loading on the substation transformers is balanced only when the load shared by each transformer in a distribution system is proportional to the capacity of that transformer. The higher the transformer capacity, the more should be the load fed by it. This loading is called ideal loading of the transformer and is calculated by multiplying the fractional capacity of the transformer with the sum of total loss and load (in MVA) on the network. Fractional capacity of a transformer is equal to the ratio between transformer capacity and the sum of capacities of all transformers in the system. For a given configuration G of the network, unbalance in transformer loading is measured by calculating the linear sum of absolute value of per unit deviation from the ideal loading for each transformer. Mathematically, unbalance in transformer loading is defined as:

$f_2(G) = \sum_j deviation_j$, where $j = \{1, 2, ...NumberOfTransformers\}$

$deviation_j$, for any j^{th} transformer is defined as the percentage deviation of transformer loading from its ideal loading IL_j, as shown below:

$$deviation_j = \{abs(LoadOnTransformer_j - IL_j)\} / IL_j, \text{ where}$$

$abs(...)$, is the absolute value function and Ideal Loading, IL_j is defined as:

$$IL_j = (TrafoCapacity_j / \sum_k TrafoCapacity_k) * TotalLossAndLoad$$

where $k = \{1, 2, ...NumberOfTransformers\}$ and

$$TotalLossAndLoad = \sum_p Load_p + \sum_q Loss_q, \text{ such that}$$

$$p = \{1, 2, ..., NumberOfBuses\} \text{ and}$$

$$q = \{1, 2, ..., NumberOfConnectedBranches\}$$

And $LoadOnTransformer_j$ is the actual load (in MVA) on the j^{th} transformer for the given configuration G of the distribution system.

3) Minimization of Voltage Deviation: Ideally, for any configuration G of the distribution system, the magnitude of voltage at each bus should be as close to 1

per unit as possible. Under normal operating conditions, the maximum deviation that is allowed in the voltage magnitude is $\pm\ 5\%$. Thus, the voltage at every bus should always lie in the range from 0.95 per unit to 1.05 per unit. Generally, in the literature, restriction of the voltage magnitude in this range has been considered as a constraint, which should always be satisfied during the optimization process. As explained in section 1, this approach is very conservative and searches only a narrow region of the search space. In this work, restriction of voltage deviation has been considered as an objective and the optimization method proposed here simultaneously optimizes it with other objectives. Mathematically, voltage deviation from 1 per unit is defined as:

$$f_3(G) = VoltageDeviation = \max\{abs(1-\min(V)), abs(1-\max(V))\},$$

where

$V = [V_1, V_2 ....., V_{n_b}]$ is a vector containing magnitude of voltage at every bus in the system.

Another objective in the multi-objective reconfiguration problem could be the minimization of maximum current through the feeder branches. We did not consider this objective in our multi-objective framework because during the search process it was observed that the maximum current in any feeder branch for any configuration did not exceed the line capacity limit for the test systems considered. So adding this objective would not have brought any additional advantage, but would have just increased the computational expense of the algorithm. However, for implementing the proposed AIS-ACO algorithm on any other system, this objective and any other objective can be included in the multi-objective framework.

Section 3 describes the approach used for solving this multi-objective optimization problem and presents the details of the AIS-ACO hybrid algorithm.

3 Nature-Inspired Computation

Nature-Inspired Computation comprises of the various computational techniques and algorithms that are based on heuristics derived from nature. These algorithms stand for a class of stochastic optimization methods that simulate the process of natural evolution. They have become immensely popular for solving optimization problems in the last decade. These stochastic methods are derivative-free and are less likely to get trapped in local minima. They could be applied on any type of search space (multimodal, discontinuous, discrete) and can be easily tailored to specifically suit a given problem. Moreover, they sample a wide region of the search space and can be hybridized with other algorithms for improving their performance. The work here combines ideas derived from two algorithms, namely Artificial Immune System and Ant Colony Optimization, belonging to the class of Nature - Inspired Computation. The theory behind these algorithms has been explained in the following sections.

3.1 Immune System

This section provides some insights into the workings of the biological immune system. The ideas presented in this section have been extracted from [32] and [33]. The Immune System is considered to be one of the most important biological mechanisms that humans possess. Recently, researchers have made significant breakthroughs in understanding the workings of the immune system [34]. The primary goal of the immune system is to keep a human healthy. It is equipped with the capability of recognizing and combating infectious foreign elements, called pathogens, which attack the human body. The pathogens contain certain molecular patterns on them, which are called antigens (Ag). In general, an antigen is any molecule that can provoke the immune system to respond. This immune system response is specific to a type of antigen. The manner in which the immune system responds is governed by the clonal selection principle [35]. The section ahead explains the clonal selection principle in detail.

3.1.1 Clonal Selection Principle

The cells that play the key role in the immune system operation are called B lymphocytes (or B cells). These cells have a molecule attached to their surface called antibody (Ab). Each B cell secretes a single type of antibody, which is relatively specific for an antigen. The primary purpose of an antibody is to recognize and bind to an antigen for marking it for elimination by other cells of the immune system. Those B cells, whose antibodies best recognize and bind the antigen, are cloned into a number of copies. This step is called proliferation or cloning. Some of these clones finally maturate into terminal (non-dividing) antibody secreting cells called plasma cells. *Plasma cells* are capable of secreting only one type of antibodies relatively specific for an antigen. The rest of the clones eventually differentiate into long-lived B memory cells. *Memory Cells* circulate through the blood and when exposed again to the same antigen, come in and convert into plasma cells. These plasma cells are capable of producing high affinity antibodies that were pre-selected for this specific antigen during its first encounter. The response of an immune system, when stimulated by an antigen for the first time, is called *primary response*. When the same antigen provokes the immune system again, then the response is called *secondary response*.

The following figure demonstrates the concepts of clonal selection principle.
In summary, the main features of the clonal selection principle are:

1) Proliferation and differentiation on stimulation of B cells with antigens.
2) Generation of new random genetic changes in the antibodies by a process called affinity maturation.

In order to further improve the affinity between the antibody and the antigen, all the memory B cells undergo a process called *affinity maturation*. Section 3.3 given ahead explains this concept in detail.

3.1.2 Affinity Maturation in Immune System

Immune system subjects some of the memory B cells' antibodies to a process called affinity maturation in an attempt to increase the antibody-antigen affinity

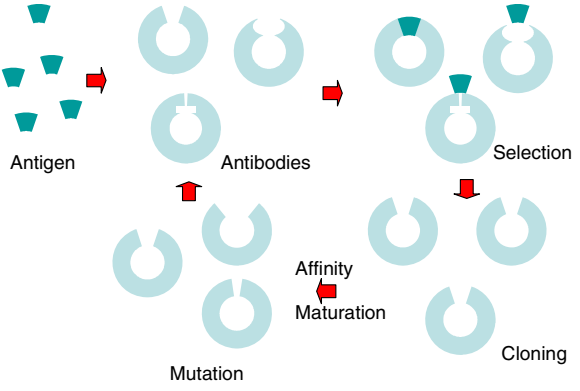


Fig. 3. Concept of the clonal selection principle

via genetic variation. This process is primarily accomplished by employing *hyper-mutation*, in which random changes are introduced into the genes responsible for the antigen-antibody interactions. Occasionally, one such change leads to an increase in the affinity of the antibody towards the antigen. These higher affinity variants are then selected to enter the pool of memory cells. The result, of having such a mechanism within the immune system, is that the responses of the immune system become faster and better.

However, the majority of the changes brought about during hyper-mutation may lead to poorer or nonfunctional antibodies. If a cell that has just picked up a useful mutation undergoes mutation at the same rate, then the possibility of losing the advantageous mutation by accumulating bad mutations is high. Thus, a selection mechanism is provided in which the regulation of hyper-mutation rate is made dependent on the affinity of the antibody. B cells with low affinity antibodies keep on undergoing hyper-mutation at a high rate, and as a result, they die off if their affinity does not improve. For cells with high affinity antibodies, the hyper-mutation becomes inactive in a gradual manner.

In addition to employing hyper-mutation on the existing memory B cells, a fraction of new B cells are added to the existing population of the memory B cells. The idea behind this is to increase the diversity of the population.

Using the ideas of clonal selection and affinity maturation described above, researchers have proposed an optimization algorithm called Artificial Immune System.

3.2 Artificial Immune System (AIS)

In recent years, several researchers have developed computational models of the immune system [36] that attempt to capture some of its features like clonal selection and affinity maturation. The algorithm derived from these models is called Artificial Immune System (AIS).

The immune system can be seen as a parallel and distributed adaptive system from an information processing perspective [37]. The ideas of the immune system have

been combined with a genetic algorithm (GA) to solve multimodal optimization problems [38]. The approach is to construct a population of antigens and a population of antibodies. Antibodies are then matched with antigens and a fitness value is assigned to each antibody based on this matching. Finally, a conventional genetic algorithm is used to replicate the antibodies that better match the antigens. Smith et al. [39] implements the ideas of immune system in such a way that fitness sharing emerges in their algorithm.

Yoo and Hajela [40] proposed the use of AIS for solving multi-objective optimization problems. Their approach uses a linear aggregating function to combine objective function and constraint information into a scalar value. This value is then used as a fitness function of a genetic algorithm. The best designs, according to this fitness value, are designated as antigens and the rest of the population is termed as antibodies. The authors then simulate an algorithm that they proposed in [41]. In [33], Coello Coello *et al.* proposed another implementation of the ideas derived from clonal selection principles for multi-objective optimization problems. Antibodies are represented by binary strings, which encode the decision variables of the problem to be solved. They do not use a separate set of antigens, but instead use the concept of Pareto dominance and feasibility of antibodies as a measure of affinity toward antigens. This information is used for determining the number of clones for antibodies. Mutation is then applied to the clones of good antibodies. The authors use a secondary population to store the non-dominated solutions/antibodies found along the search process. This secondary population is the elitist mechanism most commonly employed in the evolutionary multi-objective optimization literature.

In this work, ideas inspired from clonal selection principle and affinity maturation have been combined with ideas derived from ant colony optimization to propose an AIS-ACO hybrid algorithm for solving discrete multi-objective optimization problems. The formulation of AIS for multi-objective optimization in AIS-ACO hybrid is similar to the one proposed in [33].

Solutions to the multi-objective optimization problem in the AIS-ACO hybrid are equivalent to antibodies. There is no explicit population of antigens. The measure of antibody-antigen affinity is determined by calculating Pareto dominance among solutions/antibodies and is used for the cloning of solutions. The clones produced undergo the affinity maturation process through mutation. Instead of mutating the clones randomly (as done in clonal selection principle), mutation in our algorithm is carried out by using concepts derived from Ant Colony Optimization (ACO). The sections ahead explain the principle behind ant colony optimization and the motivation behind using concepts derived from ACO for mutation in AIS.

3.3 Ant Colony Optimization (ACO)

Ant Colony Optimization (ACO) is one of the population based meta-heuristic optimization methods for finding approximate solutions to discrete optimization problems. It has been derived from the foraging behavior or stigmergic communication – a form of indirect communication – of natural ant colonies [42-43]. ACO is

basically a solution-construction heuristic. The procedure for solution construction is based on mutual interactions among elementary agents, called artificial ants.

Any discrete optimization problem can be formulated as comprising of components derived from the problem domain. A solution to this problem is a certain combination of these components. The presence and absence of a component in a solution can be encoded by using a binary variable; where a value of 1 means that the corresponding component is present in the solution and a value of 0 means that the corresponding component is absent. For example, the components of a minimum spanning tree problem are the edges present in the graph. The solution to the minimum spanning tree problem can be formulated as a string of binary variables corresponding to the edges in the graph. A value of 1 represents the corresponding edge being connected and a value of 0 represents the corresponding edge being disconnected.

While solving a discrete optimization problem with ACO, the problem is formulated as a construction graph. The construction graph is a completely connected graph, where the nodes in the graph represent the problem components and the edges represent the transition between the components. Ants move on the construction graph to generate a solution. They lay a chemical substance, called pheromone, on the edges between the nodes of the graph, as they move along. The amount of pheromone deposited on the edges is a function of the quality of the solution that is produced. Ants' solution construction consists of transitions from node to node in a step-by-step manner. These transitions are determined by a probabilistic selection rule, based on the value of pheromones deposited on the edges between the nodes by other ants. So using the information stored in pheromone intensity, ants traverse a path in the construction graph. This path is a solution to the discrete optimization problem. Over a period of time, the path that corresponds to the optimal solution for the optimization problem gets high pheromone deposition. Any ant traversing the construction graph at this point will choose this path. In addition to pheromone intensity, some problem-specific local heuristic are also used to guide the ants through the construction graph.

ACO has been successfully applied to a large number of combinatorial optimization problems, including traveling salesman problems [42-43]; vehicle routing problems [44], [45]; and quadratic assignment problems [46]. ACO also has been applied successfully to the scheduling problems, such as single machine problems [47]; flow shop problems [48]; and graph coloring problems [49]. The section ahead presents the generic mathematical model used for ACO.

3.3.1 Mathematical Model of ACO

Let us consider that an ant has completed a partial sequence, that is, it has visited a set ξ of the nodes. Now the probability for this ant to select the node O_j, where O_j belongs to the set of the unvisited nodes, is defined as:

$$P_{ij}^{k}(t) = \frac{[\tau_{ij}(\xi)]^{\alpha} \cdot [\eta_{ij}(\xi)]^{\beta}}{\sum_{l \in N_i^k} [\tau_{il}(\xi)]^{\alpha} \cdot [\eta_{il}(\xi)]^{\beta}} \text{ where}$$

τ_{ij} is the amount of pheromone trail in between the nodes i and j

η_{ij} is the attractiveness of the solution (heuristic information) for transition between the nodes i and j

N_i is the neighborhood of node i

P_{ij} is the probability of transition between nodes i and j

And, α and β, the parameters of the ACO, determines the relative information between the pheromone information and the heuristic information.

The heuristic is usually problem specific and its design depends upon the problem formulation and availability of information on various system parameters at a given partial solution state. Ants change the pheromone values associated with the edges in the construction graph. The pheromone updating employed by ants is one of two types: local updating and global updating.

3.3.1.1 Local Updating

While building a solution, ants visit edges and change their pheromone level by applying the local updating rule before moving to the next edge. The advantage of this method is that the ant is able to propagate the information about the quality of the edge, whose pheromone is being updated, at the same instant to the ants following it. The pheromone value corresponding to an edge is updated by applying the following updating rule.

$$\tau_{ij}(\xi + 1) = (1 - \rho) \cdot \tau_{ij}(\xi)$$

3.3.1.2 Global Updating

Global updating is performed after all the ants have completed their tour. The most common strategy is to allow the best ant to lay pheromones. This scheme is known as elitist strategy in evolutionary algorithm literature. There is no dynamic information transfer in global updating. The same updating rule, as used for local updating, is used for global updating as well.

Pheromone updating marks the end of iteration. The ACO algorithm is ran until some stopping criterion is satisfied, e.g., the average quality solutions found by the ants in a generation has not significantly improved for several generations or the algorithm has been executed for a prior specified number of generations. The next section explains the inspiration behind combining ideas derived from AIS and ACO.

3.4 AIS-ACO Hybrid Algorithm

The performance of any Nature - Inspired Algorithm is extremely sensitive to the amount of exploration and exploitation it possess. Exploration is defined as the ability of an algorithm to search new unvisited parts of the search space whereas exploitation is defined as the ability of an algorithm to capitalize on the information gained from the already visited parts of the search space. If an algorithm has too much exploration then it will just keep roaming in the search space, without exploiting the possibility of finding better solutions in the vicinity of already found good solutions. This will result in slower convergence of the algorithm. On the other hand, if an algorithm has too much exploitation then it will not explore well enough to find a good solution and will get trapped in local minima. Thus, for any successful implementation of an evolutionary algorithm for solving any problem, there should be a right balance between exploration and exploitation in the algorithm.

The most common way of mutating a solution in AIS is by randomly complementing a solution component, if the solution representation is binary; or by adding/subtracting a small random number to the solution component, if the solution representation is real. The main drawback of such a mechanism is that the search process is too exploratory and lacks sufficient exploitation. Generally, to overcome this problem, some problem specific heuristics are used for directing the search. But often it is difficult to define a heuristic for a multi-objective optimization problem that can simultaneously direct the search along all objectives. Therefore, we propose a pheromone-based mutation framework in this work to improve the balance between exploration and exploitation for the AIS algorithm. Thus we call our algorithm as AIS-ACO hybrid algorithm. The main inspiration behind using pheromones for mutation is the fact that they can store information within them about the effectiveness of solutions components in improving the objective functions. Pheromones learn this information as the algorithm explores the search space during the initial stage of the search process and as the algorithm goes along, this information is used to guide the search process to improve the solution quality. Thus, by using pheromones, it is possible to learn the probabilistic distribution over the solution components, where the probability values indicate the desirability of solution components to be present in a solution. This hypothesis was experimentally tested and it was confirmed that pheromones do speed up the convergence. The results of these experiments, as shown in Section 5, demonstrate the effect of pheromones on convergence of the algorithm.

In summary, we have used the concept of pheromones derived from Ant Colony Optimization algorithm for directing the mutation process in Artificial Immune System towards optimal solutions for a discrete optimization problem. This approach has been applied to solve the multi-objective distribution system reconfiguration problem. The implementation details for the AIS-ACO hybrid for solving multi-objective distribution system reconfiguration problem are explained in section 4.

4 Solution Methodology

In this section, the solution methodology adopted for solving the multi-objective distribution system reconfiguration problem is described. As defined in section 2, the multi-objective distribution system reconfiguration problem is a search for a network topology that simultaneously minimizes real loss on the network, balances the load on the transformers and minimizes the deviation of voltage magnitude from 1 per unit, while satisfying the radiality constraint. On a graphical model of a distribution system, this problem is similar to a search for a quadratic minimum spanning tree that simultaneously optimizes all the objectives. In the literature it has been shown that quadratic minimum spanning tree (q-MST) problem is NP hard, which means that it is not possible to find an exact solution to the problem in polynomial time. In this work we propose using AIS-ACO hybrid algorithm for solving the multi-objective distribution system reconfiguration problem.

4.1 Using the AIS-ACO Hybrid Algorithm

Distribution System Reconfiguration problem is a combinatorial optimization problem and could be solved by various multi-objective evolutionary algorithms like SPEA-2 [50], NSGA-II [51] etc. Since the solution to the reconfiguration problem is a spanning tree, the use of any of these multi-objective evolutionary algorithms would require a design of a crossover mechanism that could produce a feasible spanning tree as an offspring, when applied to any two solutions from the parent population. But the design of such an operator is inherently complicated. The standard crossover techniques like one point crossover, two point crossover result in infeasible configurations that violate the radiality constraint. So for using crossover a mechanism to check the feasibility of the generated solutions needs to be designed and this would be computationally expensive for a large system. This discourages the use of these algorithms. On the other hand, Artificial Immune System uses only mutation within the feasible solution space as its primary search mechanism, so there is no need to design such an operator for it. This makes AIS the preferred choice over other multi-objective evolutionary algorithms.

Further, using the pheromones for directing search processes in AIS-ACO hybrid brings about another advantage for the multi-objective distribution system reconfiguration problem. The information learned by pheromones while solving the reconfiguration problem can also be used for restoring the distribution network under contingencies. Section 4.3 describes in detail the methodology employed for using the pheromone information for solving restoration problem under contingencies.

These observations combined together justify the use of AIS-ACO hybrid algorithm for solving multi-objective distribution system problem. Section 4.2 explains the details for implementing AIS-ACO algorithm for solving the reconfiguration problem.

4.2 Implementation Details of the Algorithm

A solution to the multi-objective distribution system reconfiguration problem is referred to as an antibody. We do not use any population for antigens. The equivalence of antibody-antigen affinity in this case is the concept of Pareto dominance among solutions as explained in section 2. This antibody-antigen affinity provides information about the solution quality. We use a separate pheromone table for representing information related to each objective. All the non-dominated solutions are then subjected to cloning, which is analogous to proliferation in Immune System. We have kept the number of clones produced for each non-dominated solution equal to the number of objectives being optimized. Thus, for multi-objective distribution system reconfiguration problem, the number of clones is equal to three, one clone for each objective. The clone of a non-dominated solution, corresponding to an objective, undergoes a pheromone-based mutation using the pheromone table associated with that objective. This step is analogous to the affinity maturation stage of Immune System. This process goes on for a certain number of iterations. Sections given ahead explain the different parts of the algorithm in detail.

4.2.1 Solution Representation

The solution for multi-objective distribution system reconfiguration problem is a network topology that optimizes all the objectives and satisfies the operational constraint. As explained in Section 2, a graphical model is used to represent a distribution system. Consider, once again, the three-feeder test network shown in Section 2.

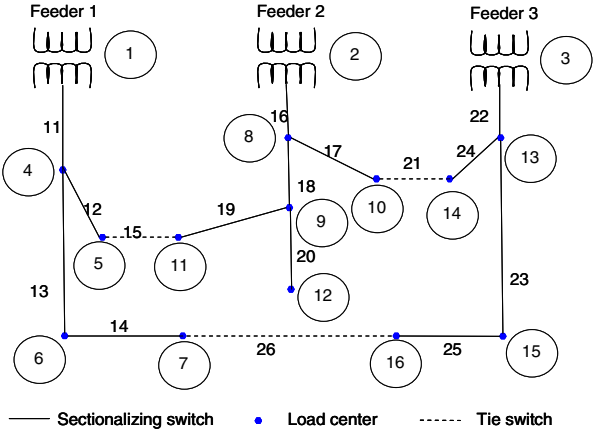


Fig. 4. A small three-feeder distribution system [23]

The figure above shows one of the feasible topology of this distribution system that satisfies the radiality constraint. For a graphical model of the network, which looks exactly the same as the figure above (with substations replaced by root

nodes as shown in Figure 2), this topology is equivalent to a spanning tree. Using the definition of a spanning tree, it can be easily seen in the figure above that some of the edges are connected and some are disconnected in a spanning tree of this system. This information could be used to represent a spanning tree in a unique way. One way could be to use the set of edges that are connected in the spanning tree and the other way could be to use the set of edges that are not connected in the spanning tree. In this work we have used the set of edges that are disconnected to represent a spanning tree. For pheromone-based mutation, this is the correct choice for solution representation. The justification behind this choice has been presented in the section on pheromone reinforcement later in this section.

So for the configuration of the system shown in the figure above, the representation of the solution is:

15	21	26

4.2.2 Initial Solution Generation

The AIS-ACO hybrid algorithm proposed in this work for multi objective distribution network reconfiguration starts with an initial population of randomly generated network topologies. The algorithm used to generate this initial set of solutions has been derived from graph theory and is called Prim's Algorithm [52]. It is a very well known algorithm and uses a set of weights attached to the edges to generate a spanning tree. In this work we do not use weights for generating a spanning tree. Instead we grow a spanning tree by randomly picking edges to connect in a partial tree. The following figure and the discussion given after it will explain the way this algorithm works.

At the start of this algorithm, the graphical model of the system is completely disconnected. Iteration begins by randomly choosing one of the root nodes, where each root node has equal probability of selection. The next step is to choose one edge at random from the set of edges emanating from this root node, which are currently disconnected. Again the probability of selection of each edge is the same. If connecting this edge in the graph doesn't create any cycle then it is

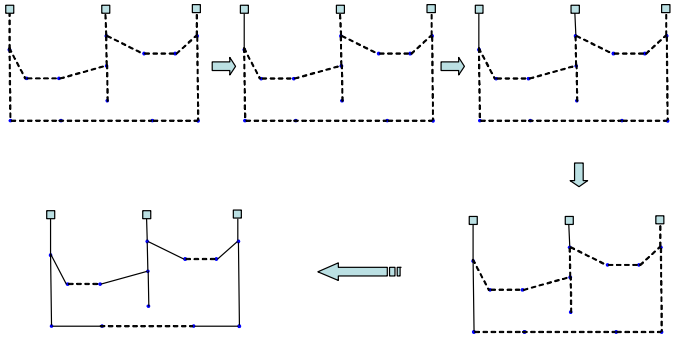


Fig. 5. Application of Prim's Algorithm to the graphical model of the three-feeder network

connected in the graph. This marks the end of iteration. The algorithm then chooses a root node again for further expansion of the graph. This process goes on till all the nodes are connected in the graph. Figure 5 shows this process. The number of times this algorithm is run is equal to the number of solutions in the initial population. For reconfiguration problem, number of initial solutions in the population is 50. The next step in the AIS-ACO algorithm is pheromone table initialization.

4.2.3 Pheromone Initialization

A separate pheromone table is initialized for each objective. It was observed experimentally that a single pheromone table was not able to store information for all the three objectives simultaneously. Hence we chose to use a separate pheromone table for each objective. For each edge in the graphical model of a network, there is a corresponding pheromone value in all pheromone tables. These pheromone values in all tables are initialized with equal value, meaning that every edge is equally likely for selection. Since a solution is represented by the set of edges that are disconnected, so the pheromone value corresponding to each edge represents the desirability of that edge to be disconnected in a solution. Figure 6 given below shows this for the three-feeder system that has 16 edges.

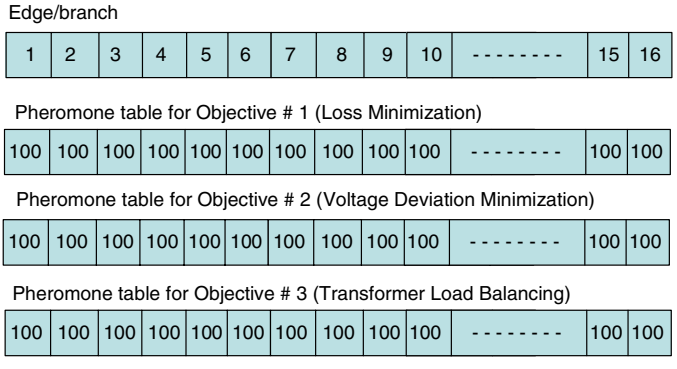


Fig. 6. Initialization of pheromone tables corresponding to three objectives

It has been shown in [53] that the performance of Ant Colony Optimization depends on the ratio of initial pheromone value and the pheromone deposition than just on the initial pheromone value. Thus pheromones can be initialized to any reasonable value (reasonable here means that the value is high enough so that it remains within necessary data type bounds for computation on a computer as the algorithm progresses) as long as the ratio of pheromone initial value and pheromone deposition can be controlled. In this work we initialized all pheromones at 100 and the ratio of initial pheromone value and pheromone deposition was set to 0.1. The next step in the AIS-ACO algorithm is fitness evaluation of each solution, which is explained in the next section. Once the fitness of all the initial solutions has been determined, the set of non-dominated solutions is extracted using the definition of Pareto Dominance, as explained in Section 2. The extracted

non-dominated solution set is then subjected to cloning and mutation. The process of cloning and mutation is explained in detail in Section 4.2.5.

4.2.4 Fitness Evaluation

The fitness of the solutions is evaluated using a load flow algorithm. For a given network configuration, a load flow algorithm calculates the currents and voltages in all parts of the network. This information is then used to compute the losses on the network, loading on the substation transformers and maximum deviation of voltage magnitude from 1 per unit. The equations representing the relationship between the bus voltages and the bus power flows are nonlinear algebraic equations. In this work, backward-forward sweep method has been used for load flow [54].

4.2.5 Cloning and Pheromone Based Mutation

Every edge in the graphical model of the network has a pheromone value attached to it in the pheromone tables. The pheromone values in the tables represent the desirability of an edge to be disconnected in the solution. For mutation, each non-dominated solution is cloned and the number of copies produced for each solution is equal to the number of objectives being optimized. For the multi-objective distribution system reconfiguration problem considered in this work the number of objectives is equal to three, so each solution here is cloned three times, one corresponding to each objective. Figure 7 given below shows this process.

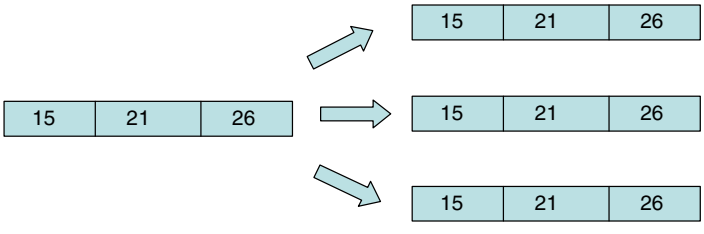


Fig. 7. Cloning process of a non-dominated solution

The first clone is mutated using pheromone table associated with loss minimization, second clone is mutated using pheromone table associated with voltage deviation and the third clone is mutated using pheromone table associated with transformer load balancing. Mutation of each clone is carried out by selecting an edge from the set of all edges for disconnection from the clone solution. This edge, which is to be disconnected in the clone, is selected by applying a roulette wheel on the pheromone table. Roulette wheel selection is a probabilistic selection mechanism where the probability of selection of an edge depends on its pheromone value. The higher the pheromone value corresponding to an edge in the table, the more is its probability of selection through roulette wheel. The following figure shows the idea behind roulette wheel selection. E#1, E#2 etc. represent the edges that are currently connected in the graph.

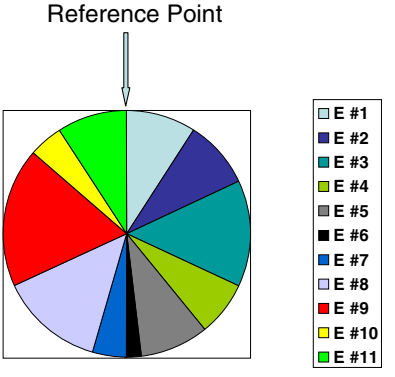


Fig. 8. Pie chart representing the effect of pheromone values on roulette wheel selection

The pheromone values in a table are normalized using the following equation.

$p_i = \frac{\tau_i}{\sum_j \tau_j}$ where $j = \{1, 2.....NumberofEdges\}$, p_i is the normalized value (also called the probability value) and τ_i is the pheromone value associated with the i^{th} edge. These normalized values are represented on a pie chart as shown in Figure 8. A random number with uniform distribution between 0 and 1 is then generated and is compared to the values in the pie chart. The edge corresponding to the sector where random number falls in the pie chart is selected for disconnecting. Once the selected edge is opened in the clone solution, a part of the spanning tree gets disconnected as shown in the Figure 9 below.

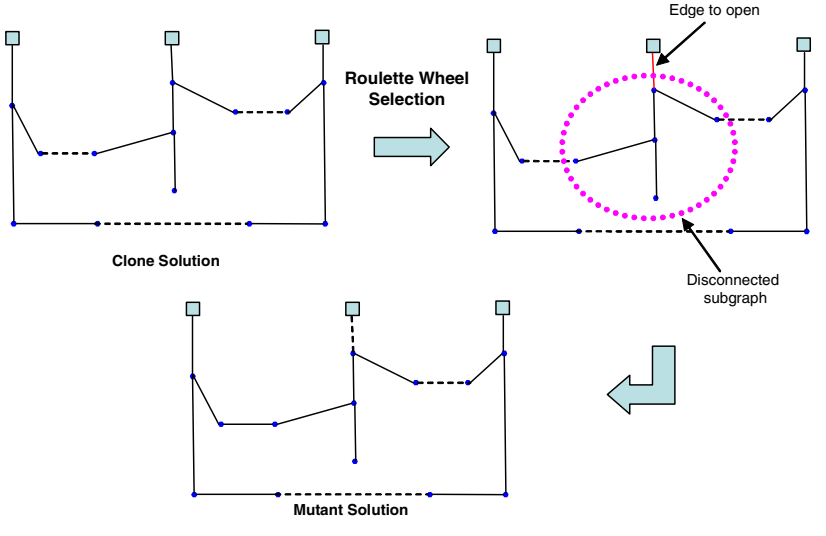


Fig. 9. Mutation process using roulette wheel selection

The next step is to find an edge that could be connected in the solution so that the disconnected part of the graph is reconnected. The choice of this edge is determined by the radiality constraint, which requires the mutant solution to be a spanning tree. All those edges are gathered that are currently open (other than the one chosen by roulette wheel) in the solution being mutated and could be connected so that the mutant solution would be a spanning tree. From this set of edges, an edge is chosen randomly with equal probability and is connected in the solution to generate the final mutant solution. For example, in Figure 9 above there are two disconnected edges that can be connected to reconnect the disconnected subgraph. Out of these one is chosen at random and final mutant solution is generated as shown. All the clones are subjected to mutation in this way and then fitness of all the mutant solutions is evaluated using the load flow algorithm. Once the fitness of all the mutants has been determined, they are combined with the parent solutions (the non-dominated set of previous iteration) and this combined set is subjected to non-dominated solution extraction. The set of non-dominated solutions is extracted for the current iteration using the definition of Pareto Dominance. The next step in the algorithm is to determine the diversity of the extracted non-dominated set.

4.2.6 Diversity Measurement

Diversity of the set of non-dominated solutions is measured in terms of the extent to which the set of non-dominated solutions is spread over in the search space. This measure can be obtained in both objective function space, also called phenotype space, and solution space, also called genotype space. However, it is preferred to calculate diversity in phenotype space, as the dimensionality of phenotype space is generally less than genotype space, so the computational effort required for diversity calculation is comparatively less. The two popular methods of measuring diversity in evolutionary multi-objective optimization literature are: 1) adaptive grid method [55] and 2) crowding distance method [51]. We have chosen crowding distance as a diversity measure in this work since it is simpler to implement and has equivalent performance when compared to adaptive grid method. The figure given below presents the basic idea behind this diversity metric. This figure represents a sample Pareto front for two objectives. The crowding distance for each solution in the front is calculated by finding the perimeter of the smallest enclosing rectangle. With three objectives, the enclosing rectangle is equivalent to a cuboid in three dimensions. Based on this value, all the Pareto solutions are sorted. Low value of this metric represents that a solution is located in a crowded area and high value represents that a solution is in less crowded area.

In this work, the diversity measure is used to trim the non-dominated solution set. We have fixed the maximum number of non-dominated solutions that will be preserved during iterations to 50. It was observed that a large number of non-dominated solutions were being found during the middle stages of the algorithm

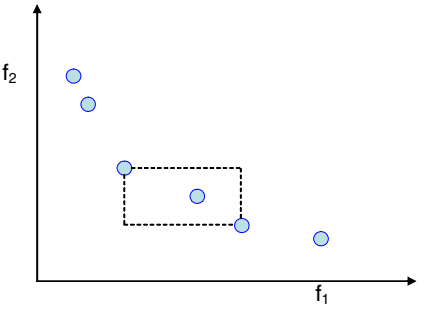


Fig. 10. The plot above shows the idea behind crowding distance metric

that were close to each in the objective space. The common belief in evolutionary multi-objective optimization community is that preserving all the closely spaced non-dominated solutions does not necessarily improve the performance of the algorithm, but instead makes the algorithm computationally expensive. So based on the position in the objective function space, non-dominated solutions are chosen for cloning in the next iteration. More solutions are selected from less populated region of the objective space and vice versa. The solutions having minimum value along each objective in the current iteration are always chosen for the next iteration. Once the non-dominated solution set has been trimmed, the next step in the algorithm is to update the pheromone tables.

4.2.7 Pheromone Reinforcement

Each pheromone table, corresponding to each objective, is updated by the solution components (edges) of the non-dominated set of solutions in all iterations. Pheromone is deposited only on those edges that show up in the non-dominated solution set. The following update rule is used for the edges of non-dominated solution set.

$$(\tau_k^j)_i = (1-\rho)*(\tau_k^j)_{i-1} + Q*(\min(F^j(G))/f^j(G))$$

where $j=\{1,2,....,\text{NumberofObjectives}\}$, $k=\{1,2,....\text{NumberofEdges}\}$, i is the iteration counter, $(\tau_k^j)_i$ is the pheromone value of the k^{th} edge in the j^{th} pheromone table at the i^{th} iteration, ρ is the evaporation rate, Q is the pheromone deposition constant, G is the graphical representation of the current solution, $f^j(G)$ is the objective value of the current solution along j^{th} objective, $F^j(G)$ is the set of objective values of all the non-dominated solutions along j^{th} objective and $\min(..)$ is the function that returns minimum value of its argument. The advantage of this pheromone reinforcement rule is that it is invariant to the scale of the objective values. When a pheromone table is updated using a non-dominated solution, the pheromone deposition factor Q is multiplied with the ratio of minimum value along that objective and the objective value of that

solution. This makes pheromone deposition independent of the absolute value of the objective and thus provides scale invariance.

The pheromone value for the edges that are not the part of non-dominated solution set just undergoes evaporation, without any deposition. The following update rule is used.

$$(\tau_k^j)_i = (1-\rho) * (\tau_k^j)_{i-1}$$

Pheromone reinforcement marks the end of iteration. The AIS-ACO algorithm is set to run for a certain number of iterations. At the end of the algorithm, the final non-dominated set is reported as the solution to multi-objective distribution system reconfiguration problem.

We can now explain the reason for choosing the set of edges that are disconnected in a spanning to represent a solution. For any spanning tree of a graph, the number of edges that are disconnected is always less than the number of edges that are connected. So if a solution is represented using the set of connected edges then pheromone will be deposited on a lot of edges. This way the difference between the maximum and minimum of the pheromone values may not be high enough, even after considerable number of iterations and pheromones may not be able to guide the roulette wheel selection well enough during the search process. But if a spanning tree is represented by the set of edges that are disconnected then this problem does not come up as the number of disconnected edges in a spanning tree is always less than the number of connected edges. So as the algorithm progresses, pheromones start to learn the probability distribution over the edges and thus are able to guide the search process. Also for a large system, representing a solution with the set of disconnected edges would be computationally less expensive compared to solution representation with the set of connected edges as the number of disconnected edges is always less compared to the number of connected edges. Thus choosing the set of disconnected edges to represent a solution for the multi-objective distribution system reconfiguration problem is the logical choice. As mentioned before in Section 4.1, pheromones in AIS-ACO hybrid algorithm also help in solving the restoration problem under contingencies. The section ahead describes this idea in detail.

4.3 Restoration Using Pheromones

Using pheromones in the AIS-ACO hybrid algorithm brings about two advantages. First, it improves the exploration-exploitation balance of algorithm and second, it provides a way to learn about relative performance of different edges in optimizing the objectives. This knowledge can be very useful in dealing with contingency situations. As mentioned in Section 1, contingency in a distribution system could be an occurrence of a fault on a feeder, on a transformer or occurrence of a sustained interruption due to failure of any component of the system. Thus, the primary objective of restoration, after the faulted section has been isolated, is to supply power to as many disrupted customers as possible and as fast as possible. This is generally achieved, before the faulted part of the network is

repaired, by changing the topology of the network such that parts without power can be fed power from unfaulted part of the network. In the AIS-ACO hybrid algorithm, pheromones provide a means by which a network can be reconfigured very quickly for supplying power to the interrupted customers. This is made possible by using the information that pheromones gather while solving reconfiguration problem under normal conditions.

Restoration problem under contingencies can be viewed as a constrained network reconfiguration problem. The feasible region of the search space for this problem contains all those network topologies that have the faulted edges open. The search for the optimum solution in the feasible region for the restoration problem, which could supply power to the interrupted customers (if possible) while staying within voltage deviation constraint and transformer and feeder capacity limit, should be as fast as possible. In literature, various methods have been proposed based on heuristics and AI based approaches [56], [57]. All these methods initially explore the feasible search space before finally converging to the optimum solution. This initial exploration for the optimal solution takes time and depending on the size of the network and location of the fault, this time delay could be too long for the utility to restore network. Thus even if the configuration obtained using these methods is optimal under contingency, but it may not be worthwhile because the fault might be cleared by that time. Moreover, quick restorations will lead to higher reliability since interruptions of less than five minutes are not counted in statistics for sustained interruptions. In the AIS-ACO hybrid algorithm, acceptable solutions to the restoration problem can be quickly generated using the information stored in the pheromones.

4.3.1 Solution Generation under Contingencies

We regenerate the network topology under contingency using the pheromones tables previously converged for network reconfiguration under normal conditions. This is a one step process and takes considerably less time compared to other approaches presented in the literature. To achieve this, the whole network is disconnected initially. The network is grown by adding edges one by one using Kruskal Algorithm [58]. Figure 10 given ahead presents the idea behind the Kruskal Algorithm. In this algorithm the choice of an edge to be connected in the network depends on the weight associated with the edge. In our problem the weights are analogous to the pheromone values corresponding to an edge. Since the values in the pheromone tables represent the desirability of an edge to be disconnected in the network topology (while here we are connecting edges one by one), so we use inverse roulette wheel to select an edge to connect in the network. In inverse roulette wheel the probabilistic selection is made on the reciprocal of the pheromone values. For solving multi-objective distribution system reconfiguration, we used a separate pheromone table for each objective. So each edge in the graph has three pheromone values associated with it, one in each table. For the restoration problem, we combined all the pheromone tables into one by multiplying together the corresponding pheromone values associated with an edge in the tables. The idea is

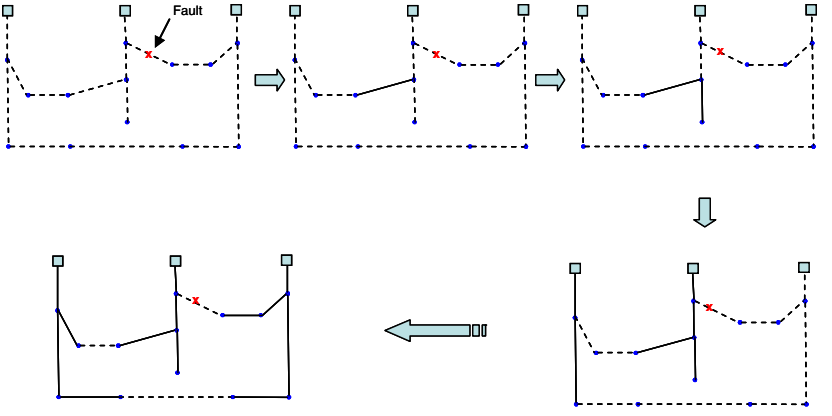


Fig. 11. Kruskal Algorithm coupled with inverse roulette wheel for generating network topology under contingencies

to use information about all the three objectives present in the pheromone tables for restoring the network.

Kruskal algorithm begins by completely disconnecting the graph. The next step is to select an edge using inverse roulette wheel on pheromones. If connecting the selected edge in the graph does not create any cycles then it is connected. This process goes on till all the nodes in the graph are not connected. The spanning tree finally obtained is the solution to the restoration problem.

The generated solution may not be the optimum under the contingency as the pheromone values were converged under normal operating conditions but nevertheless, the solution generated is locally optimal and provides a good immediate solution for restoring power. Moreover, the AIS-ACO hybrid algorithm is run in parallel to find the optimal solution under contingency. We also compared the quality of solutions produced using pheromones with the quality of solutions produced using impedance of the edges as a heuristic under contingency. The idea behind using impedance heuristic is that the real loss and voltage drop in a system are directly proportional to the impedance of the edges. So edges with high impedance should be kept disconnected.

4.4 Flowcharts for the Proposed Algorithm

This section provides flowchart for the proposed AIS-ACO hybrid algorithm and the algorithm for energy restoration using information stored in pheromones. The proposed AIS-ACO hybrid algorithm has been tested on two systems that have already been published in the literature. Section 5 ahead presents the results obtained.

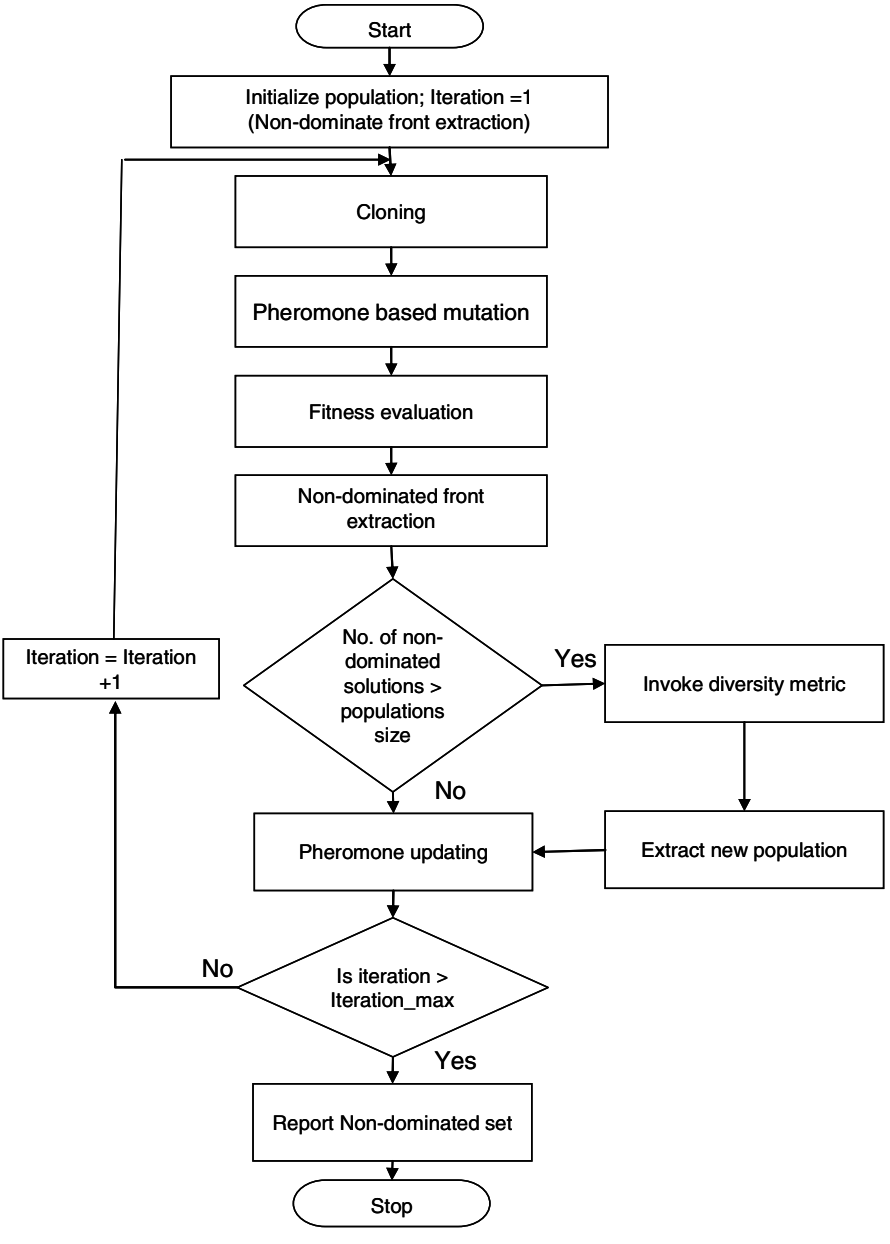


Fig. 12. Flowchart for the proposed AIS-ACO hybrid algorithm

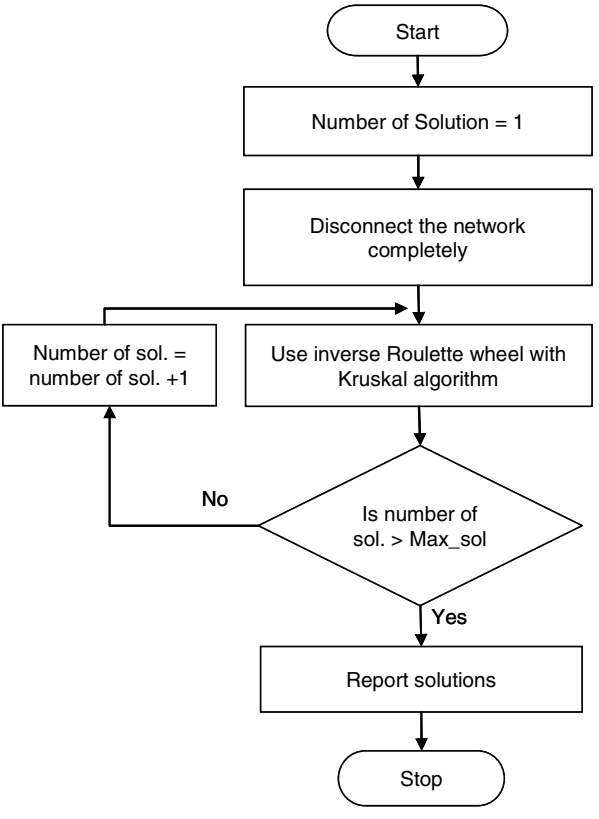


Fig. 13. Flowchart for solving Restoration problem using pheromones

5 Test Cases And Results

The first system [17] (referred to as System #1 from here on) has a total of 86 buses with three substations, 83 load buses and 96 switches and the second system [23] (referred to as System #2) has a total of 94 buses with two substations, 92 load buses and 96 switches. Data for peak loading condition was used for simulation. MATLAB software package was used for implementation and the code was run on a Pentium IV, 2.4 GHz processor.

5.1 Justification for Hybridizing AIS with ACO

The motivation behind proposing pheromone - based mutation framework for AIS is to provide direction to the search process. This hypothesis was tested experimentally. The AIS-ACO algorithm was run to optimize each objective separately. Ant colony parameter alpha, which governs the effect of pheromones on the probabilistic selection during mutation, was varied. All the other ant colony parameters

were kept constant. Even though all the parameters of ACO affect the probabilistic selection of solution components directly or indirectly, the choice of varying alpha only was based on the fact that alpha directly affects the amount of pheromone information used, as it appears as an exponent to the pheromone value during the roulette wheel selection. The values used for other ant colony parameters are Q = 10, τ_0 = 100 and ρ = 0.05. The following figures show the result obtained through these experiments.

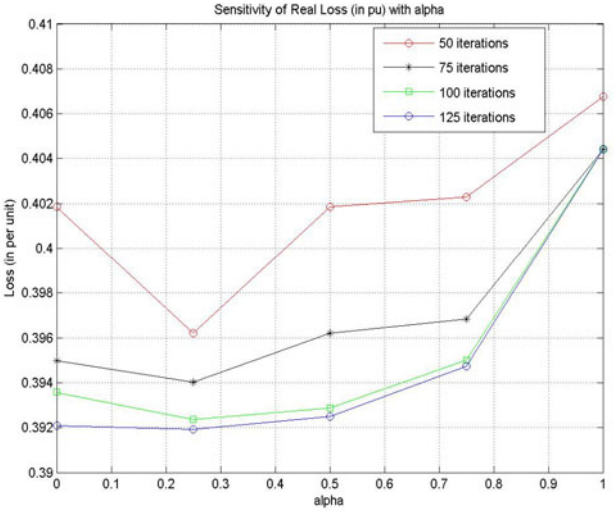


Fig. 14. Sensitivity of Real Loss (in per unit) with alpha for the System #1

The plot shown in the figure above has been obtained after averaging the results of 5 runs of the algorithm. The plot shows loss value for different values of alpha at different stages of the algorithm. The points along the x-axis represent different values of alpha and points along y-axis represent real loss values in per unit. The idea behind generating such a plot is to see the effect of pheromones on the convergence of the algorithm. Value of alpha equal to zero means the probabilistic selection in roulette wheel is independent of the value of pheromones. Hence, all the solution components (edges) have equal probability of being selected and hence the selection is basically random. As the value of alpha increases from zero, the effect of pheromones in the probabilistic selection increases as well. It is easy to observe the effect of alpha on convergence. After 50 iterations the best value found for loss is with alpha equal 0.25. This pattern continues in rest of the iterations. The conclusion drawn from this plot is that pheromones are definitely helping in finding better solutions. Thus using a pheromone based framework for mutation in AIS algorithm is justified. Same experiment was performed on remaining objectives and on the other system as well. Similar results were obtained for all the experiments. The following figures present the results obtained.

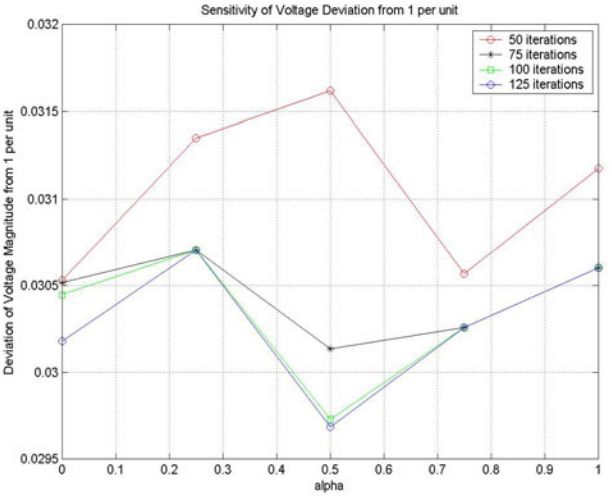


Fig. 15. Sensitivity of Deviation of Voltage Magnitude from 1 per unit with alpha for System #1

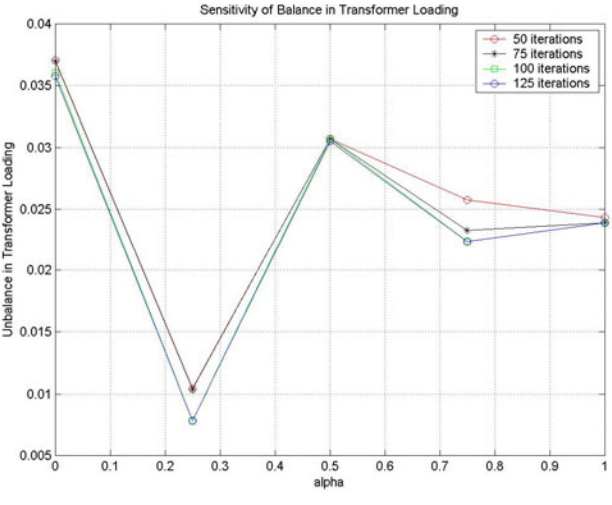


Fig. 16. Sensitivity of Unbalance in Transformer Loading with alpha for System #1

From Figure 15 and Figure 16 it can be seen that the optimum value of alpha for minimization of voltage deviation and transformer load balancing is 0.5 and 0.25 respectively. The proposed AIS-ACO hybrid algorithm was then applied to both the test systems, System # 1 and System #2, for solving the multi-objective distribution system reconfiguration problem using the methodology described in Section 4. The following section presents the results obtained for it.

5.2 AIS-ACO Hybrid Applied to the Multi-Objective Distribution System Reconfiguration Problem

Three pheromone tables were used with one corresponding to each objective. The values of ACO parameters Q = 10, τ_0 = 100 and ρ = 0.05 were the same as used for single objective optimization. The optimum values of alpha for loss minimization, voltage deviation minimization and transformer load balancing were chosen to be 0.25, 0.5 and 0.25 respectively. These values were taken from results shown in the previous section. The following figure shows convergence of the objectives function for the test system # 1.

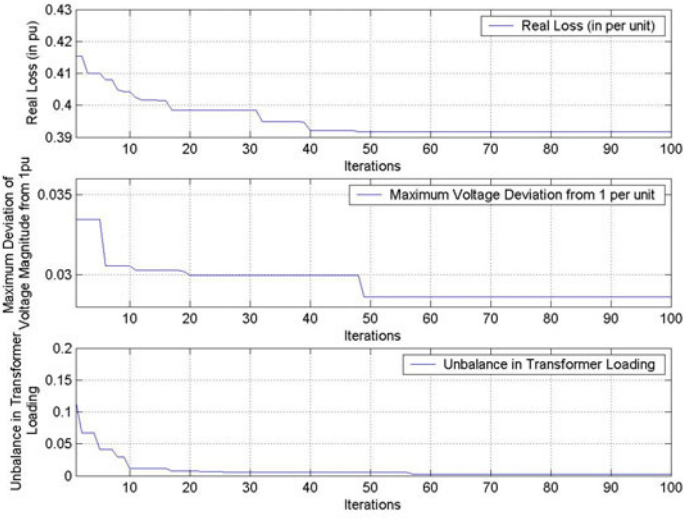


Fig. 17. Convergence Plot of the objective functions for System #1

The AIS-ACO algorithm was run for 100 iterations. This number was determined experimentally. It was observed that algorithm did not find better solutions in iterations after 100. So the AIS-ACO algorithm was stopped after 100th iteration. Figure 17 shows the improvement in each of the objective function value as the algorithm progresses. These plots show the fitness value of the best non-dominated solution obtained along each objective in a given iteration. The set of non-dominated solutions obtained at the end of the 100th iteration was reported as the final solution. The following section presents the Pareto solutions (final non-dominated set) obtained for System #1.

5.2.1 Pareto Solutions

The total number of non-dominated solutions was restricted to 50 (based on the explanation of diversity metric given in Section 4.2.6 in Section 4). Figure 18 presents a scatter plot of the Pareto solutions in three dimensions for System #1.

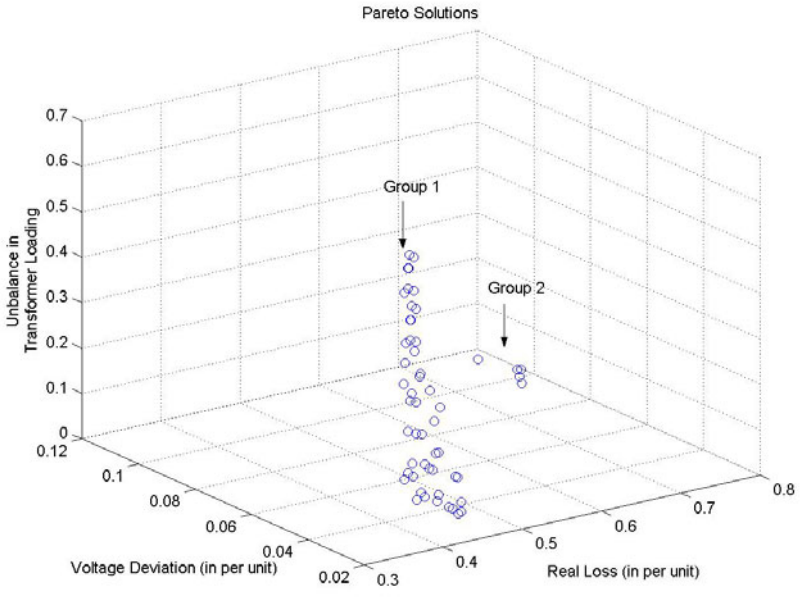


Fig. 18. Pareto solutions obtained for System #1

The Pareto solutions shown in Figure 18 are grouped into two groups in the objective space. These solutions represent the different trade-offs obtained between objective values while optimizing the objective functions simultaneously. For group 1 solutions, real loss (in per unit) ranges from 0.3919 to 0.49488; voltage deviation is less than 0.05 per unit and per unit unbalance in transformer loading varies from 0.00576 to 0.54358. So the system operator, depending on the utility priorities, can choose any solution from this group. For example, if the probability of occurrence of fault on the system is low and the utility is trying to cut the operating cost of the network then the system operator can choose a configuration that minimizes the losses the most. On the other extreme, if the main focus of the utility is to provide maximum reliability without worrying about the operational cost, then the configuration that minimizes unbalance in transformer loading to the maximum can be chosen. And if the utility wants to achieve a balance between reliability and operational cost then the system operator can choose a configuration from the rest of the solutions.

The solutions in group 2 are not dominated by group 1 solutions because they have unbalance in transformer loading lower than solutions in group 1. But losses and voltage deviation for solutions in group 2 are on the higher side. The voltage deviation is more than 0.05 per unit and loss is around 0.75 per unit. Since voltage deviation is out of the desired range, these solutions cannot be used without switching capacitors in the system.

In Figure 19 Pareto solutions have been plotted in two dimensions, with x-y axes representing different combinations of the objectives, to show the relation among various objectives. Specifically, plots of real loss versus unbalance in transformer loading, voltage deviation versus unbalance in transformer loading and real loss versus voltage deviation have been shown. It can be easily observed from the plots of real loss versus unbalance in transformer loading and voltage deviation versus unbalance in transformer loading that these objectives are in conflict with each other. This means that decreasing one objective increases the other objective. The plot of real loss versus voltage deviation shows that voltage deviation and real loss are not in direct conflict with each other. In fact they seem to be proportional to each other, which means that decreasing one decreases the other as well.

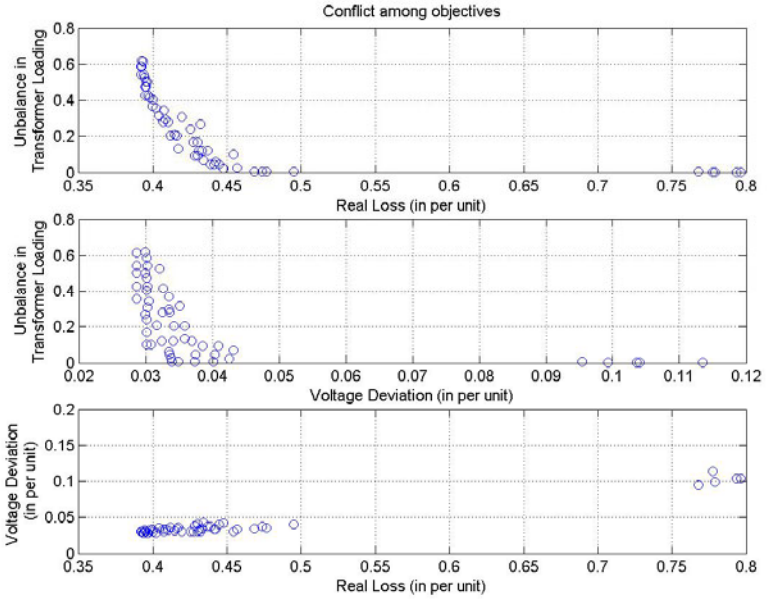


Fig. 19. Conflict among objectives for System #1

Table 1 given below presents the numerical values of the objectives for these solutions and Table 2 presents the set of switches that are open in the Pareto solutions.

L - Total Real Loss in the system (in per unit).

ΔV - Maximum of deviation of voltage magnitude from 1 per unit at all buses.

ΔT_b - Sum of percentage deviation of loads on transformers from ideal load in a configuration.

T1 – Loading on Transformer (in per unit) at Substation 1 (Bus # 1)

T2 – Loading on Transformer (in per unit) at Substation 2 (Bus # 71)

T3 – Loading on Transformer (in per unit) at Substation 3 (Bus # 53)

Table 1. Pareto Solutions found for System #1

L	ΔV	ΔT_b	T1	T2	T3
Group 1 Solutions					
0.39179	0.030265	0.54358	10.995	16.808	17.493
0.39201	0.030129	0.58888	10.653	16.808	17.836
0.39224	0.030129	0.58883	10.654	16.808	17.836
0.39253	0.029969	0.6181	10.433	16.808	18.057
0.39335	0.028633	0.61626	10.447	16.808	18.045
0.39373	0.028633	0.54314	11	16.808	17.493
0.39427	0.032139	0.52835	11.112	16.355	17.836
0.39467	0.030265	0.4296	11.857	15.953	17.493
0.39489	0.030129	0.4749	11.516	15.953	17.836
0.39512	0.030129	0.47485	11.516	15.953	17.836
0.39541	0.029969	0.50412	11.295	15.953	18.057
0.39623	0.028633	0.50228	11.31	15.953	18.045
0.39661	0.028633	0.42918	11.862	15.953	17.493
0.39771	0.032687	0.41427	11.975	15.499	17.836
0.39967	0.033507	0.3685	12.322	15.499	17.493
0.40019	0.030129	0.40326	12.06	15.42	17.836
0.40191	0.028633	0.35755	12.406	15.42	17.493
0.404	0.035152	0.31577	12.864	14.967	17.493
0.4069	0.033643	0.28193	12.98	15.42	16.93
0.4073	0.030503	0.34266	12.522	15.417	17.393
0.40812	0.033507	0.29717	12.866	15.417	17.051
0.41045	0.032498	0.28066	12.992	15.953	16.394
0.41186	0.035843	0.20533	13.562	15.953	15.827
0.41474	0.031682	0.21007	13.528	15.993	15.827
0.4159	0.034194	0.20449	13.571	15.953	15.827
0.41716	0.035843	0.13382	14.106	15.42	15.827
0.41964	0.030265	0.30821	12.79	16.808	15.761
0.42501	0.030129	0.23919	13.315	15.953	16.104
0.42702	0.030129	0.16839	13.852	15.42	16.104
0.42828	0.038585	0.092664	14.584	14.967	15.827
0.4301	0.040918	0.092476	14.589	14.967	15.827
0.43031	0.030129	0.16771	13.859	15.42	16.104
0.43105	0.032471	0.12229	14.203	15.42	15.761
0.43207	0.029969	0.26699	13.109	15.953	16.325
0.43321	0.034161	0.12123	14.213	15.42	15.757
0.4341	0.043219	0.071397	14.59	15.417	15.384

L	ΔV	ΔT_b	T1	T2	T3
0.43676	0.036953	0.12051	14.221	15.42	15.757
0.4389	0.037525	0.04626	14.784	15.42	15.198
0.44102	0.033643	0.045837	14.789	15.42	15.198
0.44225	0.033507	0.061068	14.674	15.417	15.319
0.4445	0.040344	0.045145	14.797	15.42	15.198
0.44733	0.042507	0.022886	15.256	14.967	15.198
0.45405	0.030811	0.10336	15.755	14.363	15.319
0.4544	0.030192	0.10339	15.192	14.363	15.883
0.45686	0.033814	0.027566	15.185	14.939	15.319
0.46853	0.033889	0.0078396	15.121	15.216	15.132
0.4735	0.037351	0.0073492	15.132	15.216	15.132
0.47629	0.034903	0.0070732	15.138	15.216	15.132
0.49488	0.040171	0.0057671	15.18	15.216	15.132
Group 2 Solutions					
0.76804	0.09544	0.0055037	15.337	15.417	15.384
0.77726	0.11351	0.0039605	15.358	15.417	15.384
0.77902	0.09927	0.0037895	15.362	15.417	15.384
0.79324	0.10406	0.0024109	15.393	15.417	15.384
0.79324	0.10406	0.0024109	15.393	15.417	15.384
0.7966	0.10365	0.0021176	15.401	15.417	15.384

Table 2. Open Switches corresponding to solutions in Table 1

Open Switches												
Group 1 Solutions												
13	27	31	35	39	45	49	55	57	58	66	68	72
13	27	31	35	39	45	49	55	57	58	67	68	72
13	27	31	35	39	45	49	54	57	58	67	68	72
13	27	31	35	39	45	49	55	56	58	66	68	72
13	27	31	35	39	46	48	55	57	58	66	68	72
13	27	31	35	39	46	49	55	57	58	66	68	72
14	27	31	35	39	45	49	55	57	58	67	68	72
13	27	31	35	45	49	55	57	58	66	68	72	96
13	27	31	35	45	49	55	57	58	67	68	72	96
13	27	31	35	45	49	54	57	58	67	68	72	96
13	27	31	35	45	49	55	56	58	66	68	72	96
13	27	31	35	46	48	55	57	58	66	68	72	96
13	27	31	35	46	49	55	57	58	66	68	72	96

Open Switches												
14	27	31	35	45	49	54	57	58	67	68	72	96
14	27	31	35	46	49	55	57	58	66	68	72	96
13	27	31	35	45	49	55	57	58	67	68	72	90
13	27	31	35	46	49	55	57	58	66	68	72	90
14	27	31	35	45	49	55	57	58	66	68	72	90
13	27	31	35	46	49	55	57	58	65	68	72	90
14	27	31	35	47	49	55	57	58	67	68	70	90
14	27	31	34	47	49	55	57	58	66	68	70	90
13	27	31	35	45	49	54	56	59	64	65	72	96
13	27	31	35	45	49	54	57	60	65	68	72	96
12	27	31	35	45	49	54	57	60	65	68	72	90
13	27	31	35	47	49	54	57	60	65	68	72	96
13	27	31	35	45	49	54	57	60	65	68	72	90
13	27	31	35	39	45	55	57	58	66	68	72	84
13	27	31	35	47	52	57	58	67	68	72	84	96
13	27	31	35	45	54	57	58	67	68	72	84	90
14	27	31	35	46	49	55	57	60	65	68	72	90
14	27	31	35	46	49	55	57	60	66	68	72	90
13	27	31	34	46	53	57	58	67	68	72	84	90
13	27	31	34	46	53	57	58	66	68	72	84	90
13	27	31	39	45	55	56	58	66	68	72	84	96
13	27	31	34	46	53	56	58	65	68	72	84	90
14	27	31	35	45	49	55	56	59	66	68	70	90
13	27	31	34	46	51	56	58	65	68	72	84	90
13	27	31	34	46	53	57	58	65	68	72	84	90
13	27	31	34	46	55	57	58	65	68	72	84	90
14	27	31	34	47	55	57	58	66	68	70	84	90
13	27	31	34	46	51	57	58	65	68	72	84	90
14	27	31	34	46	53	57	58	65	68	72	84	90
13	26	31	35	45	54	57	58	66	68	71	84	96
13	26	31	35	45	54	56	58	66	68	71	84	96
12	26	31	35	45	54	57	58	66	68	71	84	96
13	27	29	36	45	54	57	60	64	67	70	84	96
13	27	29	36	45	54	60	64	67	70	84	95	96
13	27	29	36	46	54	57	60	64	67	70	84	96
13	27	29	39	46	54	56	60	64	67	70	84	96
Group 2 Solutions												
2	5	14	27	31	35	49	59	67	68	70	90	95
8	14	27	31	34	44	49	55	59	65	68	70	90
2	5	11	14	27	31	35	49	57	59	68	70	90

2	5	10	14	27	31	35	49	59	68	70	90	95
2	5	10	14	27	31	34	49	59	68	70	90	95
2	5	14	27	31	35	49	57	59	65	68	70	90

In this work we have assumed that all the switches in the distribution system are automated such that all of them could be remotely opened or closed. However, in [17], from where System #1 has been taken, the authors have assumed only a set of switches to be automated. Thus direct comparison of the results obtained here with the results given in [17] was not possible.

The use of pheromones was extended to solve the restoration problem as well for both the test systems, using the methodology described in Section 4.3. The minimum loss configuration, obtained from Table 5.1, was used as the base configuration for simulating contingencies. We refer to this minimum loss configuration as 'before-contingency' configuration ahead in the section on restoration. The results of solving restoration problem using pheromones for System #1 are presented in the next section.

5.2.2 Restoration Analysis for System #1

As explained in Section 4, the pheromone tables can be reused to reconfigure the distribution network for restoring power to the interrupted part of the system under contingency. In any distribution system, the worst-case contingency could be failure of an entire substation, even though the likelihood of this happening is pretty low. We simulated this worst-case contingency on the test system and generated network topologies using previously converged pheromone tables. Also to show the effectiveness of pheromones in finding good network topologies under contingency, we generated network topologies under contingency using impedance heuristic and compared them with the topologies generated using pheromones. Since desirability of an edge staying disconnected is dependent on their impedance, the network configuration was generated by using inverse roulette wheel selection based on the impedance value of the edges (same as done with pheromone values as explained in Section 4.3 of Section 4, with impedance replacing pheromones in this case). The higher the impedance valued, the more is the probability of it being disconnected.

Ten solutions each were generated with pheromones and impedance heuristic. The execution time to generate the set of ten solutions each was around 10 seconds for both pheromones and impedance heuristic. The best three solutions out of the ten solutions have been selected. The following tables present the results obtained for System #1.

5.2.2.1 Substation Failure

NSwt – Number of Switching Operations to reach this configuration from before-contingency configuration.

Table 3. Contingency Solutions Generated Using Pheromones and Impedance Heuristic for substation failures on System #1

L	ΔV	ΔT_b	T1	T2	T3	NSwt
Using Pheromones						
Substation # 1 Failure						
0.8092	0.081126	0.39671	0	18.53	27.7	7
0.91104	0.096324	0.50248	0	29.065	17.393	9
1.0705	0.11184	0.33934	0	19.436	27.379	7
Substation # 71 Failure						
1.838	0.19283	0.51312	18.04	0	30.491	6
1.8831	0.19283	0.34274	20.149	0	28.483	8
1.959	0.2	0.41208	19.373	0	29.428	8
Substation # 53 Failure						
1.6151	0.1303	0.3329	28.014	20.019	0	9
1.9739	0.16427	0.44325	29.829	19.006	0	9
2.1227	0.17488	0.45376	30.161	19.006	0	11
Using Impedance Heuristic						
Substation # 1 Failure						
1.3422	0.12598	0.72169	0	32.267	15.155	17
2.1325	0.16359	1.3305	0	40.957	8.2329	23
1.9994	0.16492	0.83556	0	34.659	14.233	17
Substation # 71 Failure						
2.1995	0.13937	0.80646	34.617	0	14.722	24
2.4619	0.15093	0.45595	19.272	0	30.654	22
6.7306	0.38252	0.95317	15.564	0	43.907	20
Substation # 53 Failure						
3.3554	0.2656	0.10045	27.266	24.658	0	17
4.0607	0.27001	0.57012	34.376	19.125	0	21
5.0265	0.37107	0.99748	41.71	13.95	0	17

The following table presents the switches that are open corresponding to the solutions given Table 3.

Table 4. Switches open in Solutions corresponding to Table 3

Open Switches												
Using Pheromones												
Substation # 1 Failure												
1	14	27	31	35	41	42	45	55	58	67	68	72
1	14	27	31	41	42	49	55	57	58	67	68	70
1	13	27	31	35	41	42	45	49	57	67	68	70
Substation # 71 Failure												
14	27	31	35	45	49	55	57	58	67	68	85	86
14	27	31	35	45	49	55	58	60	64	68	85	86
13	27	31	35	45	54	57	58	67	68	84	85	86
Substation # 53 Failure												
13	27	31	35	45	54	58	65	68	76	77	81	96
14	27	31	35	45	55	57	65	68	76	77	81	90
14	27	31	35	45	54	57	67	68	76	77	81	90
Using Impedance Heuristic												
Substation # 1 Failure												
1	5	8	34	41	42	49	57	64	67	72	80	90
1	11	23	36	41	42	53	54	64	69	74	80	84
1	6	11	36	41	42	49	57	64	65	72	80	88
Substation # 71 Failure												
17	30	36	44	47	53	59	64	69	70	85	86	95
11	23	36	40	44	49	51	54	64	67	85	86	95
2	5	11	23	27	44	51	58	70	80	84	85	86
Substation # 53 Failure												
4	11	27	30	36	39	51	57	62	70	76	77	81
11	14	30	34	40	45	54	70	76	77	81	90	95
Open Switches												
12	25	36	40	45	53	54	57	72	76	77	81	90

5.2.2.2 Random Failure

Previous analysis considered the worst-case contingencies, that is, failure of any substation. However, contingency could occur in any part of the system. Keeping this in mind we used pheromone tables to generate solutions for some random contingencies as well. The following tables present the results obtained for System #1. NSwt – Number of Switching Operations to reach this configuration from before-contingency configuration.

Table 5. Solutions generated for failure of Switch #1 and #77 using Pheromones and Impedance Heuristic on System #1

L	ΔV	ΔT_b	T1	T2	T3	NSwt
Using Pheromones						
1.3269	0.13594	0.83319	22.3	15.873	9.2155	10
1.5926	0.14732	1.0089	19.603	20.453	7.9259	10
2.1073	0.17979	1.2035	26.233	15.417	7.4833	12
Using Impedance Heuristic						
3.7929	0.20485	1.7591	2.1239	33.062	17.716	20
4.1235	0.29932	0.6440	0	18.185	35.457	22
3.6751	0.30058	0.63533	23.12	14.838	14.68	24

The following table presents the switches that are open corresponding to the solutions given Table 5.

Table 6. Switches open in Solutions corresponding to Table 5

Open Switches												
Using Pheromones												
1	13	27	31	46	49	55	57	67	68	70	77	90
1	27	31	35	46	55	57	67	68	72	77	84	90
1	14	27	31	35	55	57	65	68	70	77	84	90
Using Impedance Heuristic												
1	5	11	20	21	39	43	51	54	64	72	77	82
1	5	9	17	34	40	43	54	57	65	69	77	88
1	11	17	29	36	40	51	54	64	70	77	84	88

The restoration problem under contingency is actually a constraint satisfaction problem. For a constraint satisfaction problem the search process tries to find as fast as possible a solution that satisfies all the constraints. More emphasis is given to the speed of the search process than to the quality of the solution. A solution which may not be the global optimum but satisfies all the constraints is an acceptable solution. For the restoration problem, there are two main constraints. The first one is to supply power to the entire disrupted network as fast as possible and the second one is to keep the voltage deviation under certain limits. For contingencies higher voltage drop is acceptable. In our analysis we considered maximum voltage drop in the system not to be more than 0.1 per unit from the nominal value of 1 per unit. Also, the number of switching operations should be kept as low as possible. Real losses on the system are not given that much priority at this point, but it is a bonus if losses are also minimized simultaneously.

The configurations generated by pheromones during substation failure and random failure, as reported in Table 5.3 and Table 5.5 presented above, seem promising when compared to the configurations generated by impedance heuristic. The maximum voltage deviation for all the configurations generated using impedance heuristic under all substation failure as well as random failure is greater than 0.1 per unit. And to reach these configurations from the 'before-contingency' configuration (minimum loss configuration with no contingency), the number of switching operations required is more compared to the contingencies generated using pheromones. While, for substation #1 failure, pheromones generated a configuration where maximum voltage deviation was less that 0.1 per unit. However, for substation # 53 failure the best configuration generated by the pheromones has voltage deviation of 0.13 per unit and for substation # 71 it is 0.19. Even though the voltage deviation is greater than 0.1 per unit for both, these configurations are still good and acceptable as the number of switching operations required to reach these configurations is less compared to the configurations generated by impedance heuristic. Moreover, the voltage deviation for these configurations can be improved to some extent by switching capacitors in the network.

Thus, pheromones can provide good and acceptable configurations almost immediately so that the system can be restored within short time. The AIS-ACO hybrid algorithm can be rerun to find the optimal configuration under contingency. The AIS-ACO algorithm was reapplied for system # 1 under substation # 1 failure to obtain the optimal solution. The result obtained is presented in the table ahead and is compared to the configurations previously generated using pheromones.

Table 7. Optimal Solution obtained for Substation #1 failure in System # 1

L	ΔV	ΔT_b	T1	T2	T3
0.57533	0.036748	0.0218	0	23.103	22.604

The voltage deviation in the optimal configuration for substation #1 as given in Table 5.7 above is 0.03 per unit. The voltage deviation for the best configuration generated using pheromones is 0.08 per unit given in Table 5.3. Comparing these two configurations we can conclude that pheromones can generate locally optimal and acceptable configurations under contingencies for fast recovery of the system. Thus the proposed AIS-ACO algorithm provides triple advantage. Firstly, it can solve multi-objective distribution system reconfiguration problem without any contingency. Secondly, the information stored in the pheromones can be used to obtain acceptable configurations for restoring power under contingencies. And thirdly, the algorithm can be applied again to find optimal configuration under contingencies.

The AIS-ACO hybrid algorithm was applied to System #2 as well. Restoration analysis, as done for system #1, was also performed for this system. The following section provides the results obtained for it.

5.3 AIS-ACO Algorithm Applied to System #2

The parameter settings of the algorithm were same as used for System #1. The following figure shows the convergence of the objective functions for System # 2.

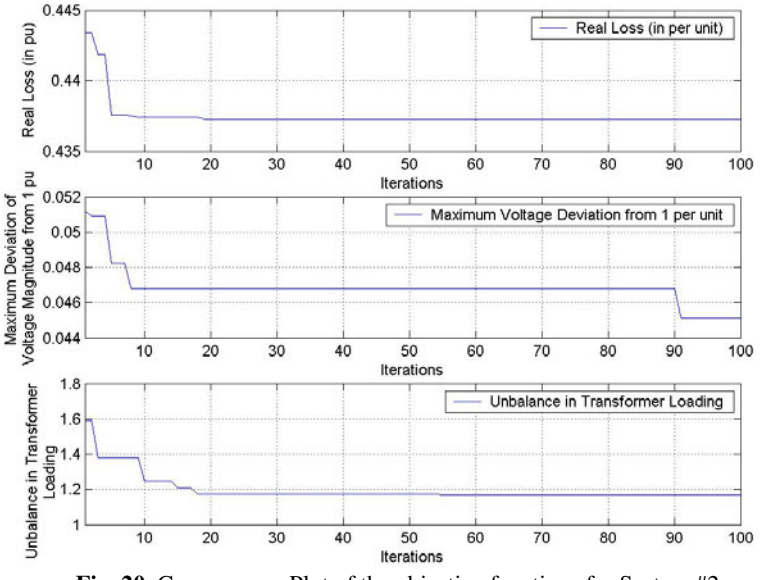


Fig. 20. Convergence Plot of the objective functions for System #2

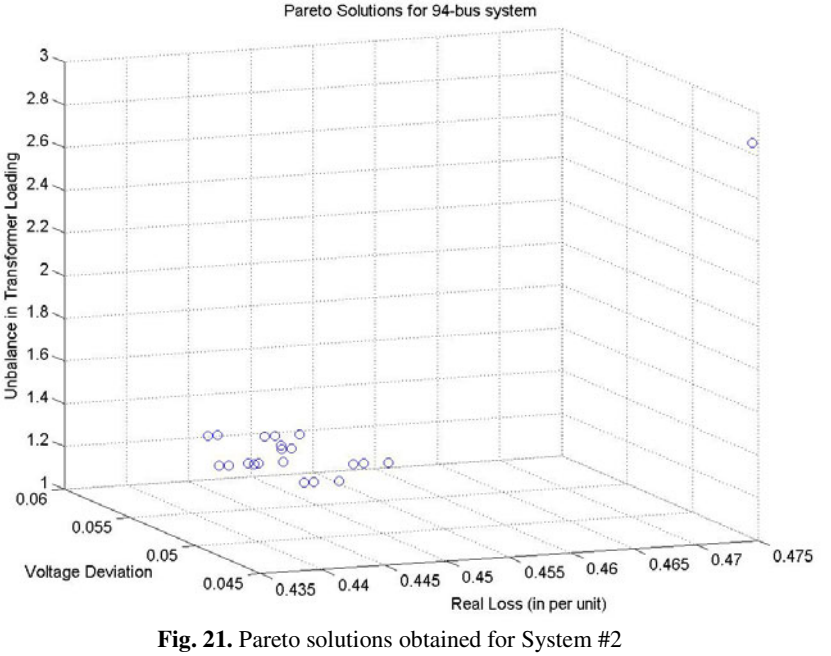


Fig. 21. Pareto solutions obtained for System #2

5.3.1 Pareto Solutions

This section presents the Pareto solutions (final non-dominated set) obtained for System #2. The total number of non-dominated solutions was restricted to 50. The following figure presents a scatter plot of the Pareto solutions in three dimensions.

Comparing Figure 18 and Figure 21, it can be observed that the Pareto solutions obtained for System #2 are located in a small region of the objective space, except for one solution which is at the corner. The reason for the clustering of solutions is related to the structure of this system. The structure is not too complex, as it does not have a lot of interconnecting lines among feeders (can be seen in Appendix A), so the search space is not very large. This means that there are not many configurations that are not dominated by other solutions. And the corner solution represents the topology that minimizes voltage deviation to the maximum but has maximum losses and poorest load distribution on the transformers.

Figure 22 below shows the conflict among objectives for System #2. Similar pattern is observed in this figure for System #2 as observed for System #1 in Figure 19.

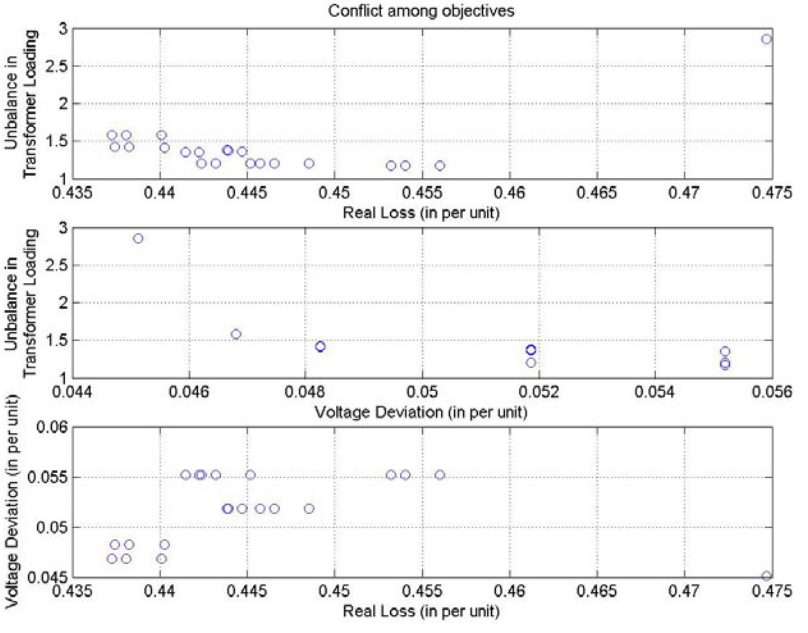


Fig. 22. Conflict among objectives for System #2

The Table 8 given ahead presents the numerical values of the objectives for these solutions and Table 9 presents the set of switches that are open in the Pareto solutions.

L - Total Real Loss in the system (in per unit).

ΔV - Maximum of deviation of voltage magnitude from 1 per unit at all buses.

ΔT_b - Sum of percentage deviation of loads on transformers from ideal load in a configuration.

A,B…..K – Loads on Transformers (in per unit) supplying power to A,B….K feeders respectively.

Table 8. Pareto Solutions found for System #2

Solution	1	2	3	4	5	6	7
L	0.4373	0.4375	0.4381	0.4383	0.4401	0.4403	0.4415
ΔV	0.0468	0.0483	0.0468	0.0483	0.0468	0.0483	0.0552
ΔT_b	1.587	1.421	1.585	1.42	1.583	1.418	1.353
A	2.834	2.834	2.834	2.834	2.834	2.834	2.603
B	2.603	2.603	2.603	2.603	2.603	2.603	2.603
C	4.545	4.545	4.545	4.545	4.545	4.545	4.545
D	2.922	2.922	2.922	2.922	3.186	3.186	2.922
E	2.924	2.924	2.699	2.699	2.442	2.442	2.924
F	2.932	2.932	3.162	3.162	3.162	3.162	2.932
G	4.214	3.95	4.214	3.95	4.214	3.95	3.76
H	2.58	2.843	2.58	2.843	2.58	2.843	3.275
I	3.267	3.267	3.267	3.267	3.267	3.267	3.267
J	2.989	2.989	2.989	2.989	2.989	2.989	2.989
K	3.252	3.252	3.252	3.252	3.252	3.252	3.252

Solution	8	9	10	11	12	13	14
L	0.4423	0.4424	0.4432	0.4439	0.4439	0.4447	0.4452
ΔV	0.0552	0.0552	0.0552	0.0519	0.0519	0.0519	0.0552
ΔT_b	1.352	1.208	1.207	1.384	1.369	1.367	1.205
A	2.603	2.834	2.834	3.487	3.487	3.487	2.834
B	2.603	2.603	2.603	2.603	2.603	2.603	2.603
C	4.545	4.545	4.545	4.545	4.545	4.545	4.545
D	2.922	2.922	2.922	2.922	2.922	2.922	3.186
E	2.699	2.924	2.699	2.924	2.924	2.699	2.442
F	3.162	2.932	3.162	2.932	2.932	3.162	3.162
G	3.76	3.529	3.529	3.598	3.574	3.574	3.529
H	3.275	3.275	3.275	2.556	2.58	2.58	3.275
I	3.267	3.267	3.267	3.267	3.267	3.267	3.267
J	2.989	2.989	2.989	2.989	2.989	2.989	2.989
K	3.252	3.252	3.252	3.252	3.252	3.252	3.252

Solution	15	16	17	18	19	20	21
L	0.4457	0.4465	0.4486	0.4532	0.454	0.456	0.4747
ΔV	0.0519	0.0519	0.0519	0.0552	0.0552	0.0552	0.0451
ΔT_b	1.205	1.203	1.201	1.174	1.173	1.171	2.858
A	3.487	3.487	3.487	3.487	3.487	3.487	2.834
B	2.603	2.603	2.603	2.603	2.603	2.603	4.383
C	4.545	4.545	4.545	4.545	4.545	4.545	5.477
D	2.922	2.922	3.186	2.922	2.922	3.186	2.922
E	2.924	2.699	2.442	2.924	2.699	2.442	2.924
F	2.932	3.162	3.162	2.932	3.162	3.162	2.932
G	3.314	3.314	3.314	2.9	2.9	2.9	4.214
H	2.843	2.843	2.843	3.275	3.275	3.275	2.58
I	3.267	3.267	3.267	3.267	3.267	3.267	0.6
J	2.989	2.989	2.989	2.989	2.989	2.989	2.989
K	3.252	3.252	3.252	3.252	3.252	3.252	3.252

Table 9. Open Switches corresponding to solutions in Table 8

Open Switches												
7	13	34	39	42	55	62	72	86	89	90	91	92
7	13	34	39	42	55	63	72	86	89	90	91	92
7	13	33	39	42	55	62	72	86	89	90	91	92
7	13	33	39	42	55	63	72	86	89	90	91	92
7	13	32	33	39	42	55	62	72	86	89	90	91
7	13	32	33	39	42	55	63	72	86	89	90	91
7	13	34	39	42	64	72	84	86	89	90	91	92
7	13	33	39	42	64	72	84	86	89	90	91	92
7	13	34	39	42	55	64	72	86	89	90	91	92
7	13	33	39	42	55	64	72	86	89	90	91	92
7	13	34	39	42	54	61	72	86	89	90	91	92
7	13	34	39	42	54	62	72	86	89	90	91	92
7	13	33	39	42	54	62	72	86	89	90	91	92
7	13	32	33	39	42	55	64	72	86	89	90	91
7	13	34	39	42	54	63	72	86	89	90	91	92
7	13	33	39	42	54	63	72	86	89	90	91	92
7	13	32	33	39	42	54	63	72	86	89	90	91
7	13	34	39	42	54	64	72	86	89	90	91	92
7	13	33	39	42	54	64	72	86	89	90	91	92
7	13	32	33	39	42	54	64	72	86	89	90	91
7	13	14	34	39	42	55	62	70	86	90	91	92

The solutions given in Table 5.8 provide the different trade-offs that can be obtained while simultaneously optimizing three objectives for system #2. A solution could be chosen depending on the utility preference. We have chosen the minimum loss configuration for the system before simulating contingencies. The results of solving restoration problem for this system using pheromones are given in the next section.

System #2 has been taken from [23]. The authors there have solved reconfiguration problem only for loss minimization. So direct comparison of results obtained using the AIS-ACO hybrid is again not possible. However, the final minimum loss configuration that the authors publish is dominated by the minimum loss configuration generated by our AIS-ACO algorithm. The result published [23] is given below:

Solution	1
L	0.4419
ΔV	0.0468
ΔT_b	1.91
A	2.834
B	2.603
C	5.127
D	3.193
E	2.924
F	2.664
G	4.214
H	2.58
I	3.267
J	2.989
K	2.684

And the switches that are open in this configuration are:

Open Switches												
7	13	34	39	41	55	62	72	83	86	89	90	92

The minimum loss configuration found by the AIS-ACO hybrid has loss equal to 0.4373 per unit. Clearly, the AIS-ACO hybrid algorithm beats the results published in [23].

5.3.2 Restoration Analysis for System #2

The following tables present the results obtained for restoration analysis on System #2.

5.3.2.1 Substation Failure

NSwt – Number of Switching Operations to reach this configuration from before-contingency configuration.

Table 10. Contingency Solutions Generated Using Pheromones and Impedance Heuristic for substation failures on System #2

Using Pheromones						
	Substation C Failure		Substation G Failure			
Solution	1	2	3	1	2	3
L	0.503	0.534	0.563	0.793	0.865	0.838
ΔV	0.051	0.054	0.064	0.138	0.145	0.145
ΔT_b	3.0036	3.1778	4.5033	4.0345	5.1393	3.8181
A	2.834	2.834	2.834	7.531	7.875	7.875
B	2.34	1.71	4.64	2.6	5.31	5.31
C	0	0	0	4.54	4.54	3.37
D	8.1	6.29	9.72	0.3	0.3	2.92
E	2.44	5.68	2.44	5.94	5.94	2.92
F	2.93	3.02	0	2.64	2.64	2.93
G	4.21	4.21	4.21	0	0	0
H	2.58	2.58	2.58	2.843	2.58	2.58
I	3.55	3.27	3.27	3.27	0.6	0.6
J	2.99	2.99	2.99	2.99	2.99	2.99
K	3.25	2.68	2.68	3.25	3.25	4.48
NSwt	5	9	13	7	7	5
Using Impedance Heuristic						
	Substation C Failure		Substation G Failure			
Solution	1	2	3	1	2	3
L	1.004	1.19	1.132	1.684	1.832	1.65
ΔV	0.127	0.137	0.152	0.194	0.207	0.216
ΔT_b	5.6569	3.8066	7.5601	10.28	12.574	9.6489
A	7.507	0.912	0.112	0.912	11.59	0.112
B	9.2	9.51	7.03	3.39	0	7.42
C	0	0	0	17	0	5.25
D	3.56	5.33	6.42	0.06	14.6	0.06
E	2.44	0	7.75	0	0	0
F	6.36	0	1.1	1.1	9.52	9.26

Table 10. (*Continued*)

G	1.87	4.18	10.7	0	0	0
H	0.824	4.709	0	10.32	0	11.64
I	0.06	7.74	0.06	3.55	0.06	0.06
J	0	1.92	1.92	1.92	0	1.56
K	4.48	2.68	1.8	0	2.68	2.68
NSwt	19	23	19	17	21	19

The following table presents the switches that are open corresponding to the solutions given Table 10.

Table 11. Switches open in Solutions corresponding to Table 10

Open Switches												
Using Pheromones												
Substation C Failure												
7	13	15	32	34	39	42	55	62	86	87	89	91
7	13	14	15	27	34	39	40	55	62	72	83	86
7	13	14	15	32	33	39	41	43	55	62	72	83
Substation G Failure												
7	13	28	34	38	42	47	63	72	86	89	90	91
7	13	28	34	38	42	47	62	70	86	89	90	91
Open Switches												
7	13	20	34	39	42	47	62	69	86	89	90	92
Using Impedance Heuristic												
Substation C Failure												
15	20	27	29	32	40	53	62	68	75	85	86	89
5	12	15	27	31	36	40	43	63	83	84	88	89
3	15	27	36	42	46	56	63	68	80	86	88	89
Substation G Failure												
5	26	30	34	40	46	47	55	77	86	87	88	89
11	14	15	31	36	39	46	47	57	62	68	73	83
3	26	27	30	36	47	55	68	76	83	86	89	95

5.3.2.2 Random Failure

The following tables present the results obtained for random contingencies on System #2.

NSwt – Number of Switching Operations to reach this configuration from before-contingency configuration.

Table 12. Solutions generated for Feeder G, J and E failure using Pheromones and Impedance Heuristic on System #2

	Using Pheromones			Using Impedance Heuristic		
Solution	1	2	3	1	2	3
L	0.843	0.844	0.929	2.021	2.045	1.711
ΔV	0.127	0.134	0.145	0.207	0.22	0.251
ΔT_b	2.2441	2.1538	2.6899	9.9019	6.4647	5.8271
A	6.995	3.487	7.875	11.59	11.57	12.02
B	5.71	5.71	5.71	0	0	9.27
C	4.54	4.54	4.54	8.06	5.13	5.13
D	5.81	5.81	2.92	0.18	8.14	6.53
E	0	0	0	0	0	0
F	3.162	3.162	6.067	18.92	1.104	2.498
G	0	0	0	0	0	0
H	3.275	6.807	2.58	0	0	0
I	3.267	3.267	3.267	0.058	10.81	0
J	0	0	0	0	0	0
K	3.25	3.25	3.25	0	2.68	2.68
NSwt	7	7	3	17	13	13

The following table presents the switches that are open corresponding to the solutions given Table 12.

Table 13. Switches open in Solutions corresponding to Table 12

Open Switches												
Using Pheromones												
7	30	33	39	42	47	64	72	73	86	89	90	91
7	30	33	39	42	47	54	72	73	86	89	90	91
7	30	39	42	47	62	72	73	86	89	90	91	92
Using Impedance Heuristic												
11	27	29	30	41	47	57	62	67	73	77	89	90
12	30	34	42	46	47	57	63	73	83	86	89	90
Open Switches												
30	34	36	41	47	55	57	66	73	83	86	89	90

The configurations generated by pheromones during substation failure and random failure for System #2 as well are better than those produced by impedance heuristic. This confirms that pheromones can provide good and acceptable configurations for restoring power very quickly.

The AIS-ACO algorithm was reapplied for System # 2 with Feeder G, J and E failure to obtain the optimal solution. The results obtained are presented in Table 5.14 and compared to the configurations generated using pheromones.

Table 14. Optimal Solution obtained for Feeder G, J and E failure in System #2

Solution	1
L	0.7809
ΔV	0.0903
ΔT_b	1.5236
A	4.821
B	4.773
C	5.477
D	5.442
E	0
F	3.518
G	0
H	5.325
I	3.267
J	0
K	3.252

The voltage deviation in the optimal configuration for Feeder G, J and E failure as given in Table 5.12 above is 0.127 per unit. The voltage deviation for the best configuration generated using pheromones is 0.09 per unit given in Table 5.14. Comparing these two configurations we can conclude that pheromones can generate locally optimal and acceptable configurations under contingencies for fast recovery of the system.

6 Conclusion

Distribution systems are the tail end of a power system. They are that part of a power system that finally delivers power to the customer premises. As the load demands of various customers change during a 24-hour period, the loading pattern of the distribution network also varies with time. At some part of the day the feeders are heavily loaded while at some other part the feeders are lightly loaded. Such variation in load is detrimental to operating conditions of the network. For example it could result in high losses on the network; voltage profile could get out of prescribed limits etc. However, with the advancement of automation in distribution

network it has become possible to automate all the switches in a network. The open/close status of these switches can now be changed in almost real time to improve operating conditions on a network. In literature, this problem is commonly referred to as Distribution System Reconfiguration. The structure of a distribution system is generally meshed but they are operated in a radial configuration. On a graphical model of a distribution network, this radial configuration is equivalent to a spanning tree.

In this work we formulated the reconfiguration problem in a multi-objective framework using the concept of Pareto Optimality. The objectives considered in the multi-objective framework were minimization of real loss, minimization of unbalance in transformer loading and minimization of deviation of voltage magnitude from 1 per unit. The set of non-dominated solutions obtained presented the trade-offs in objective values while optimizing these objectives simultaneously. Distribution system reconfiguration is an NP-hard problem, so exact method like branch and bound cannot solve it in polynomial time. In recent past, evolutionary algorithms like genetic algorithm, ant colony optimization etc. have been successfully applied to solve various classical NP-hard problems like Traveling Salesman Problem, Quadratic Assignment Problem etc. Inspired by this fact, we proposed a hybrid algorithm combining concepts from Artificial Immune System (AIS) and Ant Colony Optimization (ACO). The primary search mechanism in AIS is random mutation of genes of antibodies that affect antibody-antigen affinity. Being random, this mechanism is too exploratory and lacks sufficient exploitation of the already visited parts of the search space. To improve exploration-exploitation balance in AIS, we proposed a pheromone based mutation framework. This idea was derived from the principles of ant colony optimization. Thus we call our algorithm as AIS-ACO hybrid. Pheromones have an inherent property of storing the information gathered by the algorithm about the search space within them. This information can be used to direct the search process towards good solutions. This hypothesis was tested by experiments. It was found that pheromones do help in the search process and help the algorithm to converge faster. In addition to this, pheromones provide another advantage for solving reconfiguration problem. The information gained by pheromones while solving distribution system reconfiguration problem under normal conditions can be used to find reasonably acceptable solutions to restoration problem under contingencies. These ideas were tested on two test systems already published in the literature. The first test system has 86 buses with three substations, 83 load buses and 96 switches while the second system has 94 buses, 11 substations, 83 load buses and 94 switches. The results obtained verify the ideas and hypothesis present above.

This work proposes a generic multi-objective framework for distribution network reconfiguration. In our formulation we considered three objectives, namely, minimization of real loss, minimization of unbalance in transformer loading and minimization of voltage deviation from 1 per unit. However, this framework can include more objectives like feeder load balancing; reconfiguration for improvement of reliability indices (SAIFI, SAIDI) etc. Thus future work could include solving the reconfiguration problem with more than three objectives. Also the AIS-ACO hybrid algorithm proposed can be implemented on bigger systems to further test its efficacy.

References

1. Xu, W.: On the quadratic minimum spanning tree problem. In: Schaffer, J. (ed.) Proc. 3rd Int. Conf. Genetic Algorithms, pp. 141–148 (1995)
2. NP Hard, http://en.wikipedia.org/wiki/NP-hard
3. Goldberg, D.E.: Genetic Algorithms in search, optimization and machine learning. Addison-Wesley Longman Publishing Co., Inc., Boston
4. Dorigo, M., Maniezzo, V., Colorni, A.: Ant system: optimization by a colony of cooperating agents. IEEE Trans. Syst., Man, and Cybernetics— Part B: Cybernetics 26(1), 29–41 (1996)
5. Clerc, M., Kennedy, J.: The particle swarm: Explosion, stability and convergence in a multi-dimensional complex space. IEEE Trans. on Evolutionary Computation 6, 58–73 (2002)
6. Coello Coello, C.A.: A short tutorial on evolutionary multiobjective optimization. In: Procs. 1st Int. Conf. Evolutionary Multi-objective Optimization (2001)
7. Merlin, A., Back, H.: Search for Minimal-Loss Spanning Tree Configuration for an Urban Power Distribution System. In: Proceedings of the 5th Power System Conference (PSCC), Cambridge, pp. 1–18 (1975)
8. Baran, M.E., Wu, F.F.: Network Reconfiguration in Distribution Systems for Loss Reduction and Load Balancing. IEEE Trans. on Power Delivery 4(2), 1401–1407 (1989)
9. Aoki, K., Ichimori, T., Kanezashi, M.: Normal State Optimal Load Allocation in Distribution Systems. IEEE Trans. on Power Delivery 2(1), 147–155 (1987)
10. Shirmohammadi, D., Hong, H.W.: Reconfiguration of Electric Distribution Networks for Resistive Line Loss Reduction. IEEE Trans. on Power Delivery 4(2), 1492–1498 (1989)
11. Goswami, S.K., Basu, S.K.: A new Algorithm for the Reconfiguration of Distribution Feeders for Loss Minimization. IEEE Trans. on Power Delivery 7(3), 1484–1491 (1992)
12. McDermott, T.E., Drezga, I., Broadwater, R.P.: A heuristic non-linear constructive method for distribution system reconfiguration. IEEE Trans. Power Syst. 14(2), 478–483 (1999)
13. Morton, A.B., Mareels, I.M.Y.: An efficient brute-force solution the network reconfiguration problem. IEEE Trans. Power Del. 15(3), 996–1000 (2000)
14. Civanlar, S., Grainger, J.J., Yin, H., Lee, S.S.H.: Distribution feeder reconfiguration for loss reduction. IEEE Trans. Power Del. 142(3), 1217–1223 (1988)
15. Gomes, F.V., Carneiro Jr., S.: A new heuristic reconfiguration algorithm for large distribution systems. IEEE Trans. Power Syst. 20(3), 1373–1378 (2005)
16. Ramos, E.R., Exposito, A.G.: Path-based distribution network modeling: application to reconfiguration for loss reduction. IEEE Trans. Power Syst. 20(2), 556–564 (2005)
17. Schmidt, H.P., Kagan, N.: Fast reconfiguration of distribution systems considering loss minimization. IEEE Trans. Power Syst. 20(3), 1311–1319 (2005)
18. Momoh, J.A., Yawo, A., Arunsi, C.: Reconfiguration of Distribution Feeders for Voltage Deviation Minimization with Feasibility of ANN. In: ICPSOP 1997 Conference Proceedings, January 14-17, pp. 378–383 (1997)
19. Choi, J.H., Kim, J.C.: Network Reconfiguration at the Power System with Dispersed Generations for Loss Reduction. In: IEEE Power Engineering Society Meeting, vol. 4, pp. 2363–2367 (2000)

20. Kagan, N., de Oliveira, C.C.B.: Fuzzy decision model for the reconfiguration of distribution networks using genetic algorithms. In: Proc. 13th Power Syst. Comput. Conf., Trondheim, Norway (1999)
21. Jeon, Y.J., Kim, J.C., Kim, J.O., Shin, J.R., Lee, K.Y.: An efficient simulated annealing algorithm for network reconfiguration in large-scale distribution sytems. IEEE Trans. Power Del. 17(4), 1070–1078 (2002)
22. Delbem, A.C.B., de Carvalho, A.C.P.L.F., Bretas, N.G.: Main chain representation for evolutionary algorithms applied to distribution system reconfiguration. IEEE Trans. Power Syst. 20(1), 425–436 (2005)
23. Chiou, J.P., Chang, C.-F., Su, C.-T.: Variable scaling hybrid differential evolution for solving network reconfiguration of distribution systems. IEEE Trans. Power Syst. 20(2), 668–674 (2005)
24. Ahuja, A., Pahwa, A.: Using ant colony optimization for loss minimization in distribution networks. In: Proc. 37th North American Power Symposium, October 2005, pp. 470–474 (2005)
25. Jin, X., Zhao, J., Sun, Y., Li, K., Zhang, B.: Distribution network reconfiguration for load balancing using binary particle swarm optimization. In: Proc. Int. Conf. on Power Syst. Tech. (November 2004)
26. Roytelman, I., Melnik, V., Lee, S.S.H., Lugtu, R.L.: Multi-objective feeder reconfiguration by distribution management system. IEEE Trans. Power Syst. 11(2), 661–667 (1996)
27. Hsiao, Y.-T.: Multi-objective evolution programming method for feeder reconfiguration. IEEE Trans. Power Syst. 19(1), 594–599 (2004)
28. Das, I., Dennis, J.E.: A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. Structural Optimization 14, 63–69 (1997)
29. Delbem, A.C.B., de Carvalho, A.C.P.L.F., Bretas, N.G.: Main chain representation for evolutionary algorithms applied to distribution system reconfiguration. IEEE Trans. Power Syst. 20(1), 425–436 (2005)
30. Edgeworth, F.Y.: Mathematical Physics. P. Keagan, London (1881)
31. Pareto, V.: Cours D'Economie Politiqur, vols. I, II, F. Rouge, Lausanne (1896)
32. de Castro, L.N., Von Zuben, F.J.: Learning and Optimization using the Clonal selection principle. IEEE Trans. Evolutionary Computation 6(3), 239–251 (2002)
33. Coello Coello, C.A., Cortes, N.C.: Solving multiobjective optimization problems using an artificial immune system. Genetic Programming and Evolvable Machines 6, 163–190 (2005)
34. Janeway, C.A., Travers, P., Walport, M., Capra, J.D.: Immunobiology: The Immune System In Health and Disease, 4th edn. Garland, New York (1999)
35. Burnet, F.M.: Clonal selection and after. In: Bell, G.I., Perelson, A.S., Pimbley Jr., G.H. (eds.) Theoretical Immunology, pp. 63–85. Marcel Dekker, New York (1978)
36. Hunt, J.E., Cooke, D.E.: An adaptative, distributed learning systems based on the immune system. In: Procs. IEEE Int. Conf. on Systems, Man and Cybernetics, pp. 2494–2499 (1995)
37. Dasgupta, D. (ed.): Artificial Immune Systems and their Applications. Springer, Berlin (1999)
38. Forrest, S., Perelson, A.S.: Genetic algorithms and the immune system. In: Schwefel, H.-P., Manner, R. (eds.) PPSN 1990. LNCS, vol. 496, pp. 320–325. Springer, Heidelberg (1991)

39. Smith, R.E., Forrest, S., Perelson, A.S.: Population diversity in an immune system model: Implications for genetic search. In: Whitley, L.D. (ed.) Foundations of Genetic Algorithms, vol. 2, pp. 153–165. Morgan Kaufmann Publishers, San Mateo (1993)
40. Yoo, J., Hajela, P.: Immune network simulations in multicriterion design. Structural Optimization 18, 85–94 (1999)
41. Hajela, P., Lee, J.: Constrained genetic search via schema adaptation. An immune network solution. Structural Optimization 12, 11–15 (1996)
42. Dorigo, M., Gambardella, L.M.: Ant colonies for the traveling salesman problem. BioSystems 43, 73–81 (1997)
43. Dorigo, M., Maniezzo, V., Colorni, A.: Ant system: optimization by a colony of cooperating agents. IEEE Trans. Syst., Man, and Cybernetics— Part B: Cybernetics 26(1), 29–41 (1996)
44. Bullnheimer, B., Hartl, R.F., Strauss, C.: Applying the ant system to the vehicle routing problem. In: Voss, S., Martello, S., Osman, I.H., Roucairol, C. (eds.) Meta-Heuristics: Advances and Trends in Local Search Paradigms for Optimization, pp. 285–296. Kluwer, Dordrecht (1999)
45. Gambardella, L.M., Taillard, E., Agazzi, G.: MACS-VRPTW a multiple ant colony system for vehicle routing problems with time windows. In: Corne, D., Dorigo, M., Glover, F. (eds.) New Ideas in Optimization, pp. 63–76. McGraw-Hill, New York (1999)
46. Maniezzo, V., Colorni, A.: The ant system applied to the quadratic assignment problem. IEEE Trans. Knowledge and Data Engineering 11(5), 769–778 (1999)
47. Bauer, A., Bullnheimer, B., Hartl, R.F., Strauss, C.: Minimizing total tardiness on a single machine using ant colony optimization. Central Eur. J. Oper. Res. 8(2), 125–141 (2000)
48. Stützle, T.: An ant approach for the flow shop problem. In: Proc. 6th Eur. Congr. Intelligent Techniques and Soft Computing, Aachen, Germany, vol. 3, pp. 1560–1564 (1998)
49. Costa, D., Hertz, A.: Ants can color graphs. J. Oper. Res. Soc. 48, 295–305 (1997)
50. Zitzler, E., Thiele, L.: Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. IEEE Trans. on Evolutionary Computation 2(4), 257–272 (1999)
51. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. on Evolutionary Computations 6 (2002)
52. Prim's Algorithm, http://en.wikipedia.org/wiki/Prim's_algorithm
53. Zecchin, A.C., Simpson, A.R., Maier, H.R., Nixon, J.B.: Parametric study for an ant algorithm applied to water distribution system optimization. IEEE Trans. on Evolutionary Computation 9(2), 175–191 (2005)
54. Ranjan, R., Venkatesh, B., Das, D.: Load-Flow Algorithm of Radial Distribution Networks Incorporating Composite Load Model. International Journal of Power and Energy Systems 23(1) (2003)
55. Knowles, J., Corne, D.: Properties of an adaptive archiving algorithm for storing nondominated vectors. IEEE Trans. on Evolutionary Computation 7(2), 100–117 (2003)
56. Hsu, Y.Y., Huang, H.M., Kuo, H.C., Peng, S.K., Chang, C.W., Chang, K.J., Yu, H.S., Chow, C.E., Kuo, R.T.: Distribution system service restoration using a heuristic search approach. IEEE Trans. on Power Delivery 7, 734–740 (1992)
57. Chen, W.-H., Tsai, M.-S., Kuo, H.-L.: Distribution system restoration using the hybrid fuzzy-grey method. IEEE Trans. on Power Syst. 20(1) (February 2005)
58. Kruskal's Algorithm, http://en.wikipedia.org/wiki/Kruskal's_algorithm

Intelligent Techniques for Transmission Line Fault Classification

A.K. Pradhan

Abstract. Transmission line protection is one of the challenging functions in power system protection. The methods employed for transmission line protection in digital form include distance, current differential, phase comparison, directional comparison and travelling wave protection schemes. The transmission system protection techniques in general are involved in defining the system's state through identifying the pattern of the associated voltage and current waveforms. This means that the development a good protection scheme can be treated as a problem of pattern classification. Intelligent techniques have been successfully applied to problem of fault diagnosis, fault classification for power system protection, pattern classification/recognition in signal processing etc. Out of the artificial intelligent (AI) techniques used to solve such problems: symbolic expert system, neural network (NN) and fuzzy logic system (FLS) are the highlighted ones. The expert systems have been criticized for requiring a great effort to build (knowledge acquisition) and maintain the knowledge base. The NNs, on the other hand, offer a simple and more robust solution to pattern classification problems due to their noise suppression capacity, training power, adaptability etc. The system development time is very small in NN approach as the network parameters are mostly obtained by supervised training. On the other hand the FLS possesses the properties of being subjective and heuristic. Fuzzy-neuro technique, an integrated system enjoys the advantages of both NN and FLS. These techniques are potential candidates in pattern recognition/classification problems including that for transmission line protection.

Fault classification is essential for reliable protection of transmission lines. There are different issues of fault classification. Fault type like line-to-ground or line-to-line fault is one aspect and the other important one is fault direction estimation. Importantly the classification of fault area in a series compensated line is another challenging task. Different neural network structures and types are available in the literature for classification task. In this chapter besides monolithic structure, modular neural networks are applied for the classification of transmission faults. The classification task is divided into number of subtasks where each is accomplished by a separate network. The inputs to the various networks are provided with the current and voltage samples of the respective phase/ground. The output corresponds to the type of fault associated at that instant.

A.K. Pradhan
Indian Institute of Technology Kharagpur, India-721302

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 75 – 101.
springerlink.com

1 Introduction

A transmission system faces different types of fault during its operation. There are varieties of protection schemes available for line protection and distance relaying is common either as primary or backup. Fault classification is an essential step in distance relaying based on which the distance calculation is accomplished and overall decision is derived (including that for autoreclosing). A number of methods are available for fault classification for transmission systems and the phasor based techniques are common. Conventional techniques find limitation due to the presence of fault resistance, influence of prefault load condition and signal contamination due to noise. Neural networks have been applied for classification of transmission line faults. Different types of neural networks (supervised and unsupervised) that are being applied for fault classification include multi-layered perceptron, radial basis functional networks, neuro-fuzzy, fuzzy ARTmap, Elman, Wavelet and Probabilistic networks. These supervised networks are trained by different algorithms i.e. back propagation, orthogonal least square, Extended Kalman Filter (EKF) etc. Such neural fault classifiers can be divided into two groups based on the input vectors one dealing with the sampled values of current and voltage and the other using phasors. The former approach provides very fast classification whereas the latter would be more reliable. A neural network structure available for fault classification of transmission line considers all the three phase voltage and current samples for input vector and the network has 11 outputs representing for all types of faults. Another NN approach has been shown using voltage and current samples of three phases as input for 5 outputs showing the state of three phases, ground and direction. In this network the output is assigned to 1 if the corresponding phase is involved with the fault else it is 0 and for forward direction (say) 1 is assigned otherwise for reverse fault it will be 0. Such an approach reduces the number of outputs.

Distributed representation and strong learning capabilities are the major features of neural network. Fuzzy logic systems (FLS), on the other hand, base their decisions on inputs in the form of linguistic variables derived from membership functions. The variables are then matched with the specific linguistic IF-THEN rules and the response of each rule is obtained through fuzzy implication. NN has the shortcoming of implicit knowledge representation, whereas, FLS is subjective and heuristic. The major limitations of FLS are the lack of a general systematic procedure for rule learning and tuning and determining the best shape of membership functions. As NN and FLS have different advantages and drawbacks, it is quite reasonable to consider the possibility of integrating the two parade gms into the same system in order to benefit from both of them. One such approach is integrating the learning capabilities of neural network to the robustness of fuzzy logic systems in the sense that fuzzy logic concepts are embedded in the network structure. It also provides a natural framework for combining both numerical information in the form of input/output pairs and linguistic information in the form of IF-THEN rules in a uniform fashion. Similar to the neural network for fault classification, with fuzzy-neuro systems is shown using voltage and current sampled values as input and 4 outputs corresponding to three phases and one ground.

The series capacitor in a compensated line introduces further problem due to reduction in effective reactance and functioning of metal oxide varistor (MOV) during fault. In a series compensated transmission line load current level may be of the order of fault current toward boundary of a zone which will lead to wrong classification and malfunction of the conventional relay. Further in the presence of a series capacitor at the midpoint the current level may be of the same order at two different fault points of the transmission line (one in front of the capacitor and the other behind it) for the same type of fault. This will result in more complexity while locating the fault point on a transmission system. Therefore, in a series compensated line a more reliable fault classification approach is necessary for identifying the type of fault and the fault section; behind or in front of the capacitor. Wavelet transform which is powerful in capturing information on different frequency bands is suitably applied for series compensated line fault classification in conjunction with fuzzy logic.

Majority of the faults in a transmission line are of line-to-ground fault type and the voltage and current waveforms of that faulty phase are best representative of its state. In case of a monolithic structure (conventional neural network classifiers) all the signals being taken as input, signals of sound phases will be redundant at that situation and this introduces unwanted complexity to classification. Therefore, the available techniques with single neural network (monolithic structure) have limited performance capability.

Modularity substantially reduces the complexity of the problem by decoupling the unnecessary information sharing between various units. Especially for neural networks the problem of over-fitting, under-fitting and the problems associated with large training data are almost eliminated. The modular concept for neural network has been borrowed from the principle of *divide and conquer*. In this approach any task is divided into number of possible subtasks where each one is accomplished by an individual neural network. Finally all network outputs are integrated to achieve the overall task. Obviously the approach has the advantages of simplicity, accuracy, less training sets and time, with easier interpretation. For the transmission line fault classification, there is separate network for each phase and ground and the individual network takes corresponding line current and line voltage samples only. Each network indicates the state of that phase (or ground); involved with the fault (or not). Overall the 10-class (LG, LL, LLG, and LLL faults with all phases) classification task is divided into four 2-class problems and instead of a single network four networks have been designed. Finally all network outputs are integrated to achieve the overall task.

In this chapter different transmission line fault classifiers are presented which includes conventional neural network structure to recently introduced modular neural network and wavelet-fuzzy combined approach.

2 Radial Basis Function Neural Network

Multilayered feedforward neural network (MFNN) is widely used model for signal classifications. This is because the structure of the MFNN allows it to generate internal representations directed to classify the input regions that may be either

disjointed or intersecting. The hidden layer nodes in the MFNN can generate hyperplanes to partition the input space into various regions and the output nodes can select and combine the regions that belong to the same class. There are varieties of MFNN including adaline, multi-layered back propagation NN, Radial basis function NN (RBFNN), wavelet network etc. RBFNN has the advantages of fast training and clustering action of its hidden layer which provides a way to good initialization for the training of output layer. The second point provides the benefit of avoiding local minima. In this section RBFNN based transmission line fault classification is described.

The RBFNN consists of an input layer consisting of source nodes and a hidden layer of enough dimension for nonlinear mapping. The output layer supplies the required response of the network to the activation patterns applied to the input layer. The nodes within each layer are fully connected to the previous layer as shown in the Fig.1. The input variables are each assigned to a node in the input layer and pass directly to the hidden layer without weights. The hidden nodes or units contain the radial basis functions (RBFs).

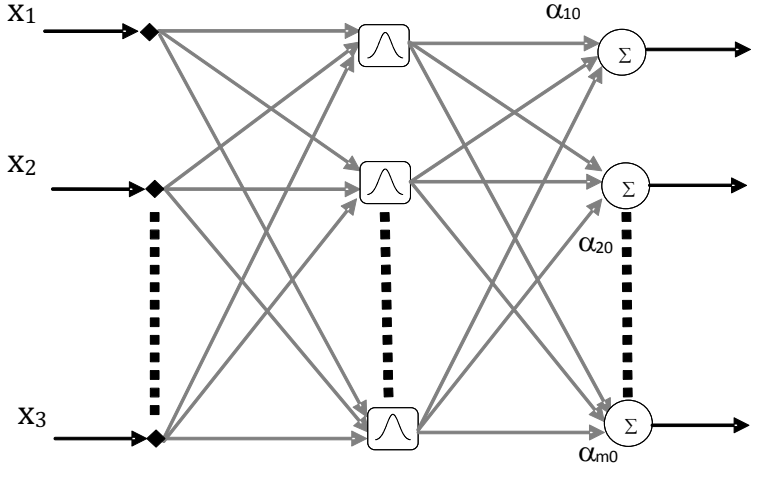


Fig. 1. The Radial Basis Neural Network Structure

Each hidden unit in the network has two parameters called a center (μ), and a width (σ) associated with it. The activation function of the hidden units is Gaussian function in this case. The response of each hidden unit is scaled by its connecting weights (α's) to the output units and then summed to produce the overall network output. The overall network output is therefore

$$\hat{y}(n) = f(x_n) = \alpha_{m0} + \sum_{k=1}^{K} \alpha_{mk} \phi_k(x_n) \tag{1}$$

For each input x_n, n represents the time index, K=number of hidden units, α_{mk}= connecting weight of the kth hidden unit to output layer, α_{mo}=bias term, m is the number of output.

The value of $\phi_k(x_n)$ is given by

$$\phi_k(x_n) = \exp(-\frac{1}{\sigma_k^2}\|x_n - \mu_k\|^2) \quad (2)$$

Where μ_k is the centre vector for the *kth* hidden unit and σ_k is the width of the Gaussian function; $\|\ \|$ denotes the Euclidean norm.

The network starts with no hidden unit and as training sets are received, new units are added. The network parameters are updated by the EKF training method. The pruning strategy removes those hidden units which contribute insignificantly to the overall network output consistently. A sequential growth and the pruning strategy provides a compact network structure, therefore, this form of RBF neural network is termed as minimal RBF neural network (MRBFNN).

Application of Minimal RBF Neural Network to Fault Classification

To obtain an accurate classification in distance relaying scheme using MRBFNN for any power system (say, for a system as in Fig.2) following steps are to be carried out.

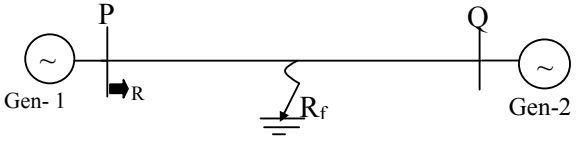


Fig. 2. The transmission system

Feature Extraction

Neural Networks learn the desired input-output mapping using the training examples without looking for an exact mathematical model. Once an appropriate neural network is designed, the weights of the network represents for the non-linearity of the desired mapping between the inputs and outputs. One of the important issue in pattern recognition is to reduce the size of the input data; known as feature extraction. In distance protection scheme, few parameters of the identified characteristics must be extracted to represent the state of the transmission line. This preprocessing stage can significantly reduce the size of the neural classifier which also improves the performance and speed of the design process.

The voltage and current data are the common information in the power systems. The fault voltage and current signals are often noisy. In addition, when a fault occurs on a transmission line, voltage and current signals develop a decaying dc component along with different harmonics whose magnitudes depend on many factors that are random in nature. Thus the input data must be preprocessed before being used by the network. For a protection application there are two options to feed the voltage and current information to the network; (i) a sampled data window or (ii) extracted fundamental and/or non-fundamental components. Out of these two the former approach requires more number of inputs to the network and vis-à-vis the classification task seems to be more complex in comparison to the latter. This implies that the former approach would require a larger structure but

faster in operation. Again from dependability and security aspects of relaying the trip decision should be as accurate as possible, therefore, the fundamental components will be a secured choice for the network input in the classification task without much compromising the speed of operation. The inherent benefit in this choice is that the conventional relaying scheme extracts the fundamental components from noisy data during fault conditions to take the decision of trip. Further without waiting till the filter output converges to estimate fundamental component, an alternative approach would be to consider the filter output during its transient condition as input to the neural network. This approach is supported by the fact that a neural network provides a special kind of flexible filtering.

Network Input Selection

During a power system fault it is expected that the current level of the faulty phase would be too high, however this may not happen in high fault resistance situation towards the line end. Again the distance estimation, which reflects the impedance up to the fault point, will be a tough task if only current information is fed to the network. A series of study reveals that both current and voltage information of each phase (in combination) reflect the state of that phase more transparently. On the occurrence of a fault the current of the faulty phase goes up and the voltage goes down. At the initial stage of the fault the signals contain high frequency components and current waveform is modulated by decaying dc component.

When only one phase is involved with fault, it is obvious that ground is also associated. However for the case of double line and double line-to-ground faults the fundamental magnitudes information is not sufficient to distinguish. Therefore zero sequence component of current, as conventionally used, could be suitably opted for ground detection in the event of a fault. Further for estimation of fault distance only magnitudes of phasors information will not be sufficient and all fault types do not contain zero sequence component. To solve this negative sequence substitutes the zero sequence component of current. Finally, from the above discussions it is concluded that the input vectors to the networks should constitute of all fundamental current and voltage magnitudes (as estimated by a filter in its transient mode), sequence component of current and the system frequency estimated at the relaying point. Kalman filtering techniques or other suitable methods can be utilized to estimate the above inputs where frequency is obtained by the model-1 as mentioned. The current signals being distorted by decaying dc component an extended Kalman filter model can be chosen. The sequence components are then derived from these current phasors. A Kalman filter is selected for fundamental voltage estimation where harmonics are also modeled. A sampling rate of 1-3 kHz would be typical choice for modern relaying applications. Though the phasor values can be directly fed to the neural structure, to reduce the complexity preprocessing like sequence components is advantageous. Another important feature for distance relaying is system-operating-frequency which not only modulates the impedance of the line and affects the level of fault current and voltage at the relaying point.

The Distance protection Scheme

On a three phase power system, there are 10 distinct types of possible fault. As the equations that govern the relationship between currents and voltages at the relay

location are different for each of the fault type, a conventional scheme uses number of distance relays energized by different pair of voltage and current signals. This implies that the current and voltage patterns of 3 phases are different for different types of fault. Therefore in applying pattern recognition techniques to transmission line protection a single distance estimator for all types of fault will be inappropriate. However it is expected that there is a close relationship between waveforms of faulty phases in case of each category of fault (phase to ground or phase to phase or double phase to ground faults).

On the basis of the above discussion, four RBFNNs, one network each for the four categories of fault (LG, LL, LLG, LLL), are shown in the distance protection scheme. This will ensure convergence while training the networks for the above purpose. Further, a fault classification unit consisting of a RBFNN is designed to select the phases and ground involved with the fault. Based on the output of the classifier the control unit fires the proper RBFNN in the fault location block. The diagram showing the major blocks of such a scheme is presented in Fig. 3 .

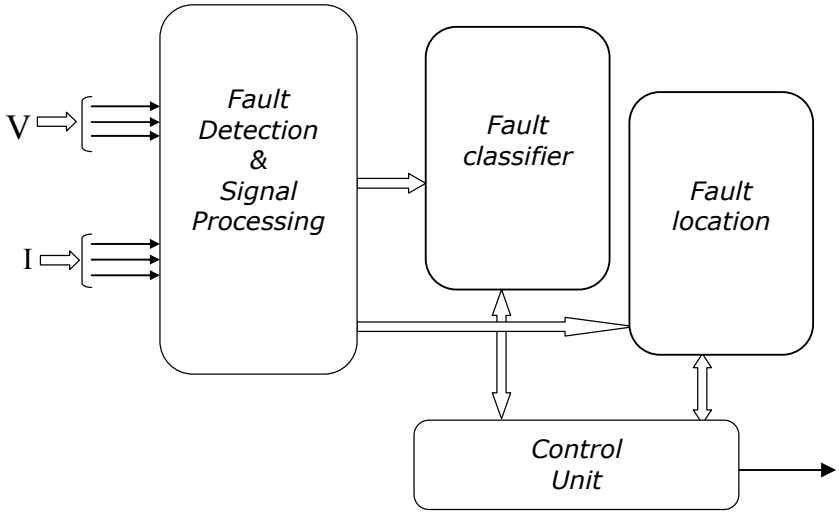


Fig. 3. The distance relaying scheme

The task of the distance protection is to command the trip signal only when a fault occurs within 80% of the transmission line length in the forward direction. The direction away from the busbar and towards the overhead line is considered as forward direction. This direction needs to be identified so that the relay does not respond for faults occurring at the bus or the line behind the relay. A fault detection unit registers a fault in the line within around 3ms (three samples in an 1kHz sampling rate) and triggers the fault classification unit to select the faulty phases/ground.

Design of MRBFNN classifier

The distance protection scheme needs a fault type identification unit to select the proper fault location unit. A MRBFNN is designed for the purpose. The algorithm

provides a sequential growth of the network and vis-a-vis the pruning strategy facilitate for a compact structure. The design of the RBFNN consists of two parts; training and testing of the network and is outlined below.

Training

In order to learn a specific relationship, a neural network must be exposed to proper data sets while training. These sets consisting of input and associated output vectors should cover the whole range of expected situations. Since fault is detected by a separate unit the training and testing of the network are accomplished using data during faulty condition of the network. In this case the output of the network should indicate the type of fault. In the present case, the MRBFNN has four outputs representing 'a', 'b', 'c' phases and 'g' ground (Fig.4). During training these outputs are assigned either 1 or 0 considering whether the fault is involved with that phase/ground or not. For example,

'a'	'b'	'c'	'g'	
1	0	0	1	-------- a-phase to ground fault (LG)
1	1	0	0	-------- a-phase to b-phase fault (LL)
0	1	1	1	-------- bc-phases to ground fault (LLG)

Owing to the lack of field data, normally simulations using the EMTDC software package are used to generate the sample data required for training and testing of the network.

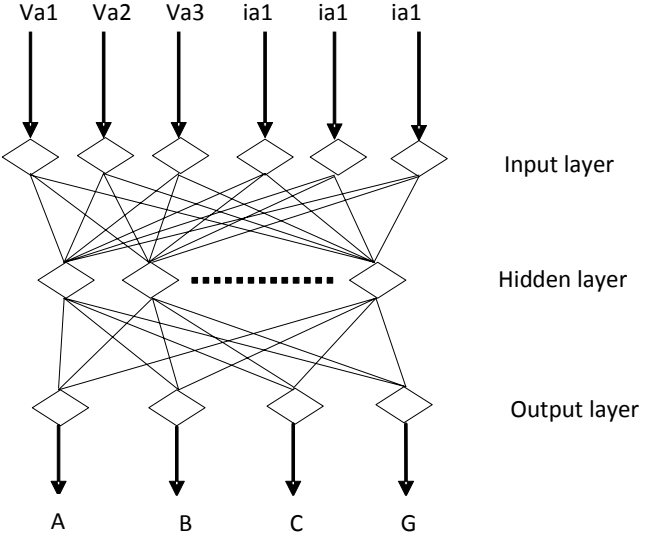


Fig. 4. Structure for the neural network classifier

Testing

The performance of the network is tested using voltage and current data of the power system during various types of shunt faults at different situations of the system. The objective of the testing is to evaluate the speed, generalization and noise immunity of the RBFNN classifier. For this, tests were carried out to see the capability of the MRBFNN in classifying the faults with different prefault condition, frequency, fault path resistance, location and inception angle. Thus a fast, accurate and robust fault classifier can be designed for any transmission network using MRBFNN.

3 Fuzzy Neural Network

A neural network which can perform pattern matching task has a large number of highly interconnected processing nodes that has the ability to learn and generalize from training patterns. Distributed representation and strong learning capabilities are the major features of neural network. Fuzzy logic systems (FLS), on the other hand, base their decisions on inputs in the form of linguistic variables derived from membership functions. The variables are then matched with the specific linguistic IF-THEN rules and the response of each rule is obtained through fuzzy implication.

NN has the shortcoming of implicit knowledge representation, whereas, FLS is subjective and heuristic. The major limitations of FLS are the lack of a general systematic procedure for rule learning and tuning, and determining the best shape of membership functions. As NN and FLS have different advantages and drawbacks, it is reasonable to consider the possibility of integrating the two paradigms into the same system in order to benefit from both of them. One such approach is integrating the learning capabilities of neural network to the robustness of fuzzy logic systems in the sense that fuzzy logic concepts are embedded in the network structure. It also provides a natural framework for combining both numerical information in the form of input/output pairs and linguistic information in the form of IF-THEN rules in a uniform fashion.

In this section a simple neural network is used to implement a fuzzy-rule-based classifier of a power system from input/output data. The FNN model can be viewed either as a fuzzy system, a neural network or a fuzzy-neural system. The structure is seen in neural viewpoint for training and fuzzy viewpoint is utilized to gain inside into the system and to simplify the mode. In this strategy the number of rules needed is determined by the data itself and consequently a smaller number of rules is produced. The network is trained using back propagation algorithm. To have a compact structure, a pruning strategy eliminates the redundant rules and fuzzification neurons.

The input and output vectors can be same as considered in radial basis function neural network in the previous section. It is to be noted that the fuzzification process in a fuzzy-neuro technique provides a special kind of flexible filtering, faster measuring algorithms that speed up the operation of protective relays could be used.

The Fuzzy Neural Network

Fuzzy neural networks can be divided into two main categories. One group of neural networks for fuzzy reasoning uses fuzzy weights in the neural network. In the second group, the input data are fuzzified in the first or second layer, but the neural network weights are not fuzzy. The latter approach is considered for application to transmission line protection as a powerful AI technique. Here, the structure can be viewed either in fuzzy or neural network sense and the weights and parameters of the network are tuned using neural network learning technique.

Fig.5 shows the architecture of the fuzzy neural network, comprising input, fuzzification, inference and defuzzification layers. Further the network can be visualized as consisting of N inputs, with N neurons in the input layer and R rules, with R neurons in the inference layer. There are NxR neurons in the fuzzification layer and K neurons for output layer. The signal propagation and basic function in each layer of the FNN is introduced in the following.

The input layer consists of x_i , i = 1,2, …N, along with unity. Each neuron in the fuzzification layer represents a fuzzy membership function for one of the input variables. The activation function used in this layer is $f(\text{net}_{ij}) = \exp(-\left|net_{ij}\right|^{l_{ij}})$ and the input to these neurons $\text{net}_{ij} = w_{ij1}\, x_i + w_{ij0}$, with w_{ij1} and w_{ij0} being the connecting weights between input layer and fuzzification layer.

Thus, the output of the fuzzification layer becomes

$$\mu_{ij}(x_i) = \exp(-\left|w_{ij1}x_i + w_{ij0}\right|^{l_{ij}}) \tag{3}$$

Where μ_{ij} is the value of fuzzy membership function of the i^{th} input variable corresponding to the j^{th} rule. The connections between fuzzification and inference layers have unity weights (shown in the figure as **I**).

Each node j in the inference layer is denoted by Π, which multiplies the input signals and the output of the node becomes the result of product. Therefore, the output of the layer becomes

$$\rho_j(x_1, x_2 .... x_N) = \prod_i^N \mu_{ij}(x_i) \tag{4}$$

With v_{jk} being the output action strength of the k^{th} output associated with the j^{th} rule and utilizing weighted sum defuzzification, the network output becomes

$$\begin{aligned} O_k(x_1, x_2, ...x_N) &= \sum_j^R v_{jk}\, \rho_j(x_1, x_2, ...x_N) \\ &= \sum_j^R v_{jk} \prod_i^N \exp(-\left|w_{ij1}x_i + w_{ij0}\right|^{l_{ij}}) \end{aligned} \tag{5}$$

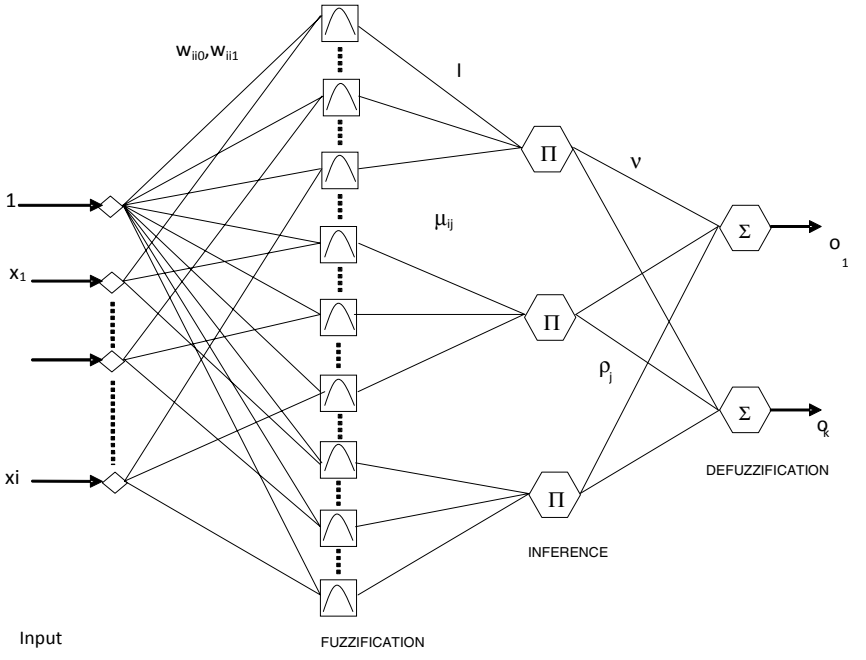


Fig. 5. The fuzzy neural architecture

Fuzzy view of the FNN

Inputs to the fuzzification layer are the process variables used to define fuzzy operating regions. Each of these variables is transformed into several fuzzy sets in the fuzzyfication layer. Through changing the weights and parameters, appropriately shaped membership function at different positions can be obtained. Each neuron in rule layer corresponds to a fuzzy operating region of the particular classification objective. Its inputs are the fuzzy acts which determine the corresponding operating region. Its output is the product of its inputs and is the membership function of the corresponding fuzzy operating region. There are no weights to be estimated in this layer.

Neural Training of the FNN

Back propagation (BP), by far the most used training algorithm in NN, can be exploited to update parameters of the fuzzy neural network.

The error function E of the network be,

$$E = \frac{1}{2}\sum_{k}(t_k - o_k)^2 \tag{6}$$

Where t_k is the desired output in the k^{th} output node. The design can be accomplished by any other alternatives like EKF. The weights between the inference and output layers are updated using the BP algorithm. During training, the number of rules is increased from 1 till a satisfactory performance of the network is found.

The Pruning Strategy

Even if a single fuzzy membership function is near zero over its input range, the output of the corresponding rule becomes close to zero. As this rule does not contribute to the network performance, the rule should be pruned. Further, with multiple inference a fuzzy membership function having close to unity over its input range contributes negligibly to the network output. This neuron can be also eliminated without hampering the network performance. By removing these redundant rules and neurons form the structure a compact form can be achieved. To implement this technique, we run the trained network with the same training sets once and see the outputs in inference and fuzzification layers. In the event of such situations exist, corresponding neurons are pruned and then network performance is studied.

The FNN Fault-Type Classifier

Using the fundamental or sequence components of voltage and current signals available at the relaying point a distance relaying algorithm derives the trip decision for faults on a transmission line. The presence of different sequence components of current reflect the type of fault, but to derive a correct fault classification out of it is a complex and time consuming task. The sequence components of current (if any) decrease as the fault point moves away from the relay. The ratios of different sequence components of current may provide still a better solution. This is evident from Fig.6 where a line to ground fault type is considered for a transmission system to compute the ratio of zero to positive sequence component of current. The ratio decreases with distance and with the inclusion of 100 Ω fault path resistance the curve becomes a different one. This complex situation therefore, demands for technique like fuzzy neural network. Fundamental frequency components of current and voltage can be taken into account along with different sequence components for a distance protection scheme as they provide more information.

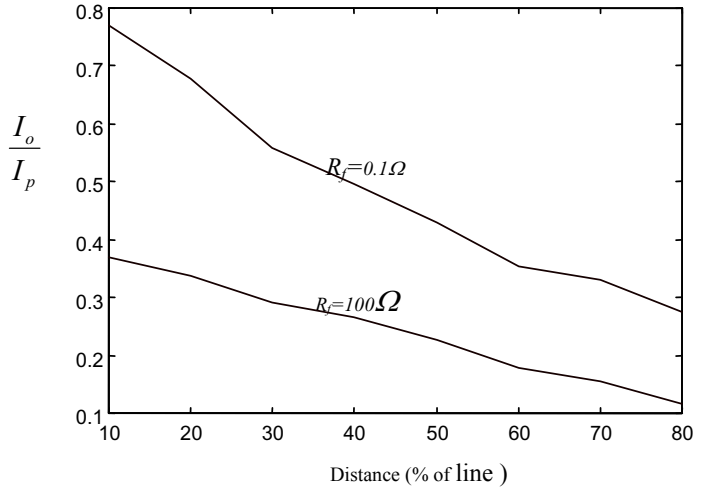


Fig. 6. Effect of fault resistance on sequence components (line to ground fault)

Training and Testing of FNN Classifier

For classification of fault normalized values of post fault peaks of fundamental components of voltages and currents of the three phases, system frequency and zero sequence current are considered as input vector for the FNN. These current and voltage peaks are estimated from sampled values of signals by different estimators. The FNN consists of four outputs representing *'a', 'b', 'c'* (phases) and '*g'* (ground). During training these outputs are assigned '1' or '0' considering whether the fault is involved with the phase/ground or not. The training sets include data for different fault locations, for different fault inception angles, source impedances, fault resistances, different prefault conditions and different operating frequencies of the system. The number of rules is increased from 1 during the training process till satisfactory response of the network is derived. With BP algorithm the network is trained and finally pruning strategy is applied. The network was tested using voltage and current data of the power system during various types of shunt faults at different locations, inception angles, fault resistances, operating frequencies and prefault conditions of the system. This classification approach takes a particular phase to be 'involved' with fault if its corresponding value is greater than a threshold value else it categorizes the phase/ground to be 'unassociated' with the fault. To show the convergence speed of the network, one example cases is presented in Fig.7 for ag-type fault on a sample to sample basis. The faulty phase/ground curves reach nearly to 1.0 within half a cycle after fault inception and the network consistently provides the accurate result (with 1 kHz sampling rate).

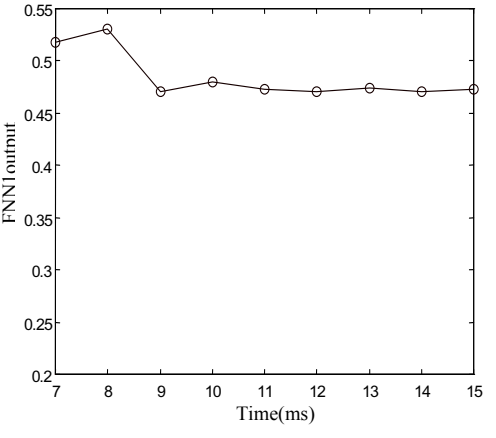


Fig. 7. Convergence loci of the FNN1 for 'ag' fault

Discussion

To view the FNN classifier in the fuzzy sense the membership functions of a particular input in rule one of the fault classification solution to the system of

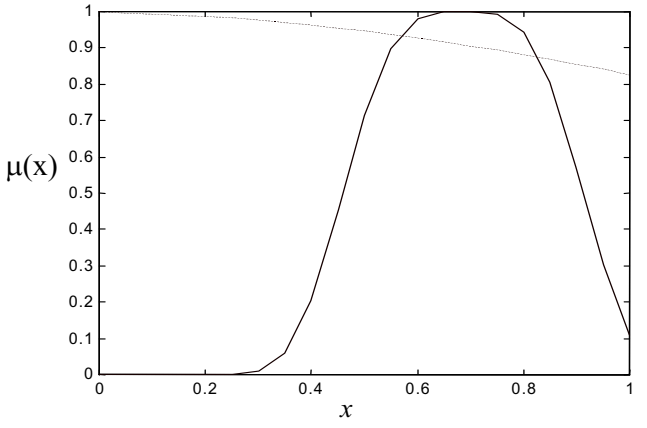


Fig. 8. Fuzzy membership function of the 6th input in rule one before training ……after training————

Fig.2, before training and after training, are plotted in Fig.8. The input to the fuzzification neuron is modulated by the w_{ij1} and w_{ij0} terms which are initialized with random values. As these two parameters are updated during training the fuzzy membership function varies. Also during training the l_{ij} value which affects the nature of the curve is tuned with training from an initial value of 2; for

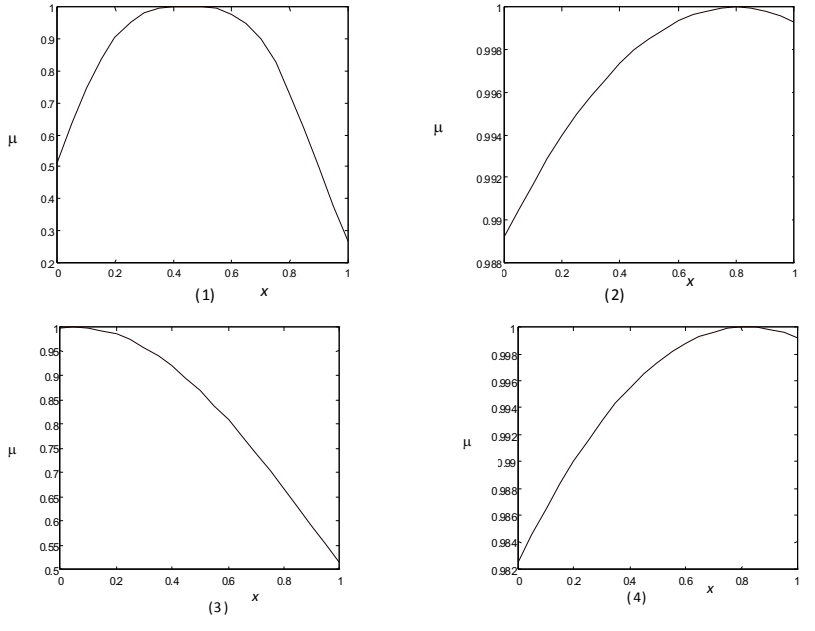


Fig. 9. Fuzzy membership functions of one input variable in 4 fuzzification neurons

example, the l =2 before training is modified to l =3.6 after training in the fault classification task. Another example of the FNN classifier showing the different membership functions of the 4^{th} input in the 8 fuzzy-neurons (as an example) are depicted in Fig.9. Interestingly, it is to be noted that the learning algorithm captures membership functions based on evidence contained in the data, not on any *a priori* choice about its form or shape.

4 Fault Classification in Series Compensated Network

Series compensations are widely used today for better utilization of transmission systems. Such compensation possesses a capacitor which is protected by a metal oxide varistor (MOV) and an air-gap arrangement. During a power system fault, current level in the circuit increases and voltage across the capacitor may exceed the limit when the MOV functions, followed by firing of the air-gap. The non-linear behavior of a series capacitor arrangement during fault-situations affects the current and voltage signals and thus, creates problems to relay functionality.

Identification and classification of faults on a transmission line are essential for relaying decision and auto-reclosing requirements. A conventional approach classifies the fault based on fundamental frequency only. However, in a series compensated line load current level may be of the order of fault current toward boundary of a zone which will lead to wrong classification and malfunction of the relay. Further in the presence of a series capacitor at the midpoint the current level may be of the same order at two different fault points of the transmission line (one in front of the capacitor and the other behind it) for the same type of fault. This will result in more complexity while locating the fault point on a transmission system. Therefore, in a series compensated line a more reliable fault classification approach is necessary for identifying the type of fault and the fault section; behind or in front of the capacitor.

Wavelet transform is highlighted for obvious advantages of better time-frequency localization and providing richer problem-specific information from sensor signals. Fault signals in a power system contain fundamental, high frequency and decaying dc components. Therefore, wavelet transform can be a good candidate for providing more vital information to on-line classification objectives. Fuzzy rule-based systems are applied to classification problems where non-fuzzy inputs are to be assigned to one of a given set of classes. In this section, fault classification of a transmission line possessing a series capacitor at midpoint is carried out using fuzzy logic systems that consider wavelet coefficients of current signal as inputs. Meyer wavelet, which has very fast decay wavelet function, is considered to extract the required information from current signals for the fault classification. The maximum values of wavelet coefficients at three levels are fed to a fuzzy logic system (FLS) designed for fault type classification. Considering only two variables derived from the wavelet coefficients the faulty section; in front of or behind the series capacitor is derived by another fuzzy logic system. Unlike other approaches based on sequence components the new method identifies the faulty phase/ground from current information of that phase/ground only. The approach uses the different frequency components of the fault current signal for successful classification.

Two classification problems are attempted for protection of the line (i) classification of the fault type and (ii) identification of fault section; F_1 or F_2 section of the line as shown in Fig. 10. For a typical transmission line the fault current level depends on the fault distance. But it is evident from the figures that section identification (F_1 and F_2) from amplitude of fault current is not possible. However, for accurate relay decision and exact fault location fault section should be identified correctly.

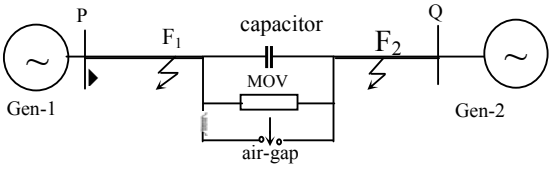


Fig. 10. The series compensated line

The wavelet fuzzy integrated approach

The series compensated transmission line fault classification task can be divided into two parts; fault section identification and fault type classification. The flow diagram of the fault classification technique is shown in the Fig.11. The classification tasks are accomplished by integrating wavelet transform with fuzzy logic system. Meyer wavelet in discrete form is applied to individual current signal at the relaying location and the corresponding maximum value of wavelet coefficients (details and approximation) at level-1 and 2 are calculated. The three indices of such maximum values distinctly possess the characteristics of the involvement of that particular phase/ ground. To detect involvement of ground with the fault, the sampled current signal considered is sum of the sampled values of all three-phase currents at each instant. For on-line implementation, a fuzzy classifier is incorporated which further improves the performance. The FLS is of min-max type where the final decision is based on the max operation. Similarly identification of the fault section (left/right of the series compensation) is done by another FLS designed. This FLS provides output by considering wavelet transform of a faulty phase current signal. The input to the FLS is different from earlier one and only two inputs, derived from the maximum values of the wavelet transform coefficients, are used for the purpose.

Feature Extraction

The knowledge of fundamental component information of current signals may not be sufficient for a secured fault classification. Discrete wavelet transform of current signal, which contains information on high frequency components also, can be of utility for the task. Meyer wavelet, in particular is a frequency band-limited function whose Fourier transform is smooth, unlike the Fourier transform of many other wavelets. This smoothness provides a much faster decay in time. The feature of Meyer wavelet results in providing high frequency information of fault signals with distinctive efficiency for classification purpose. The lower scale wavelet coefficients reflect on the high frequency components of the signal,

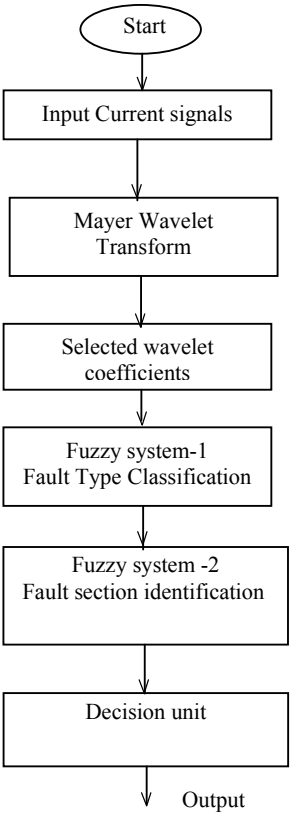


Fig. 11. Flow diagram of the wavelet transform based fault classification scheme

during fault and that to pre-fault situations. This specialty of Meyer wavelet transform is utilized for fault classification of a series compensated line. Different fault situations were created for the transmission system and only current signals are used in the wavelet transform. Some of the wavelet coefficients are provided here which reflect on the involvement of fault. For fault type identification the normalized peak values of wavelet coefficients of details at level-1, details at level-2, and approximation at level-2 etc are considered as the fuzzy input variables. Generally, the normalization is accomplished by dividing the selected wavelet coefficients with their corresponding maximum values as obtained from the simulation study of the system. For numerous fault conditions these coefficients are found suitable as input features for fault type classification. For trip decision and fault location requirement, identifying the fault behind and in front of the capacitor is essential for which these input variables are not suitable. This is because the MOV acts non-linearly during fault situations.

For fault section identification two ratios; (peak of detail at level-1 to peak of approximation at level-1) and (peak of detail at level-2 to peak of approximation at level-1) are considered to provide the required information and their normalized

values are the inputs to the second FLS. The normalization is carried out by dividing the ratio with its corresponding maximum value as obtained for a system. These values are higher for fault behind the capacitor than for faults in front of the capacitor.

Fuzzy Classifier

After obtaining the normalized peak values of the coefficients of wavelet transform of fault signals, a fuzzy logic based expert system is used to classify the type of fault. In place of FLS neural network or fuzzy-neuro technique can be suitably designed, For knowledge on simple FLS a design process is provided here. As power system data are uncertain and fault classification is a pattern recognition task, the fuzzy expert system is well suited to this kind of problem. The fuzzy expert systems, in this work, are designed using knowledge about different kinds of faults occurring both left and right of the series capacitor. FLS considered in this section is described below.

Fuzzy IF-THEN rules for a C-class pattern classification problem with n attributes can be written as

$$\text{Rule } R_j : \text{If } x_1 \text{ is } A_{j1} \text{ and } \text{-----} x_n \text{ is } A_{jn} \text{ then class } C_j,\ j=1, 2, ., N$$

Where $\mathbf{x}=(x_1, \ldots..x_n)$ n- dimensional pattern vector, A_{ji} is antecedent linguistic value such as small and large (i =1, 2, ….,n), C_j is consequent class (one of the given C classes) and N being the number of fuzzy IF-THEN rules. With grid-type fuzzy partition (triangular in this case), the antecedent part of each fuzzy IF-THEN rule is specified by a combination of linguistic values and total number of possible combinations is K^n when each attribute x_I has K linguistic values. The compatibility grade ($\mu_j(\mathbf{x})$) of the fuzzy IF-THEN rule R_j is found out by the minimum operation.

$$\mu_j(\mathbf{x}) = \min\{\mu_{j1}(\mathbf{x}_1), \mu_{j2}(\mathbf{x}_2), \ldots..\mu_{jn}(\mathbf{x}_n)\} \quad (7)$$

where $\mu_{ji}(\mathbf{x}_i)$ is the membership function of the antecedent linguistic value A_{ji}.

Without considering certainty grade (strength of a rule) a pattern is classified by the single winner rule R_{j*} defined by

$$\mu_{j*}(\mathbf{x}) = \max\{\mu_j(\mathbf{x}) : j=1,2,\ldots.N\} \quad (8)$$

Two fuzzy logic systems designed for the two classification tasks are mentioned below.

A typical fuzzy membership curve for first FLS is shown in Fig.11 and corresponding FAM rule is provided in Table1.

Table 1. FAM rules for phase selection

Rule 1: If input-1 =L and input2 =L and input3 =L then no *fault*.
:
Rule 14: If input 1=H and input2 =H and input3 = H then *fault*.

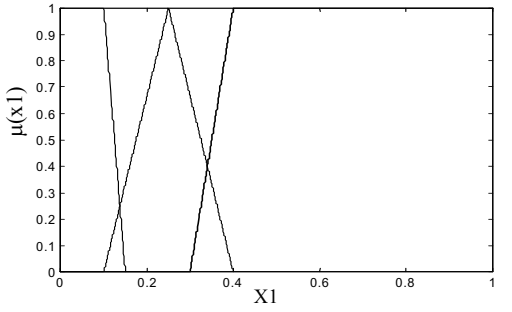


Fig. 12. Fuzzy membership function of input-1 for FLS-1 (phase selection)

Similarly for Fault Section Identification the second FLS is designed and corresponding membership curve and FAM rules are provided in Fig 12 and Table 2 respectively.

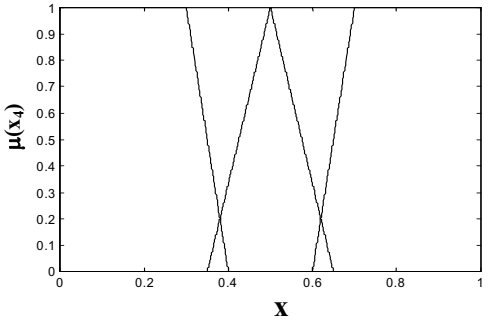


Fig. 13. Fuzzy membership functions for second FLS

Table 2. FAM rules for faulted section identification

Rule 1: If input4 =L andinput5 =L then fault is on the *left,*
:
Rule 9: If input 4 =H and input 5 =H then fault is on the *right*.

5 Modular Neural Network Classifier

The earlier neural network model presented is of monolithic in structure and leads to a complex proposition handling all possible fault types. Modular neural concept

is a recent proposition which simplifies the complex task by *divide and conquer* principle. The formulation of modularity concept is based on the fact that brain is organized into relatively independent functional units which work in parallel to accomplish a given task. In a similar fashion a complex task is first divided into different subtasks which are solved by a number of specialized computational units (modules). Subsequently, the overall solution is obtained by combining the result of each module. In such a process each module is independent of each other, responds to specific subtask and each one has a simpler architecture. Further such a module can respond to a given input vector faster than a complex monolithic system. A conventionally used monolithic neural network structure is shown in Fig.14 where whole task is carried out with one network. One type of modular neural network structure is shown in Fig.15 where a number of networks combine to solve a task.

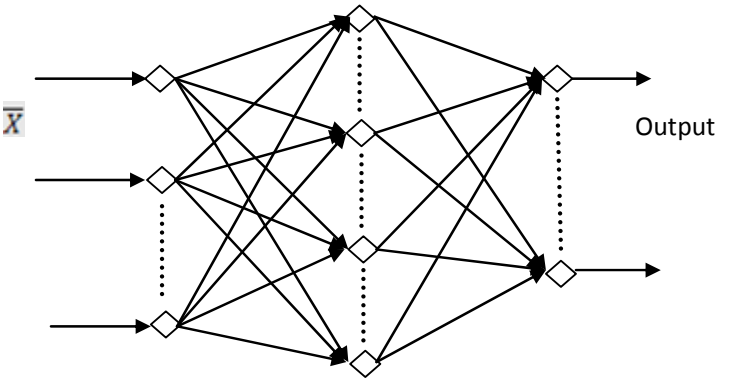


Fig. 14. The monolithic neural network structure

Modular Neural Fault Classifier Structure

In this section, different structures of modular neural networks are considered for fault classification of transmission systems. Two types of classification problem are considered to be solved by the approach. The first one is for a simple transmission system where different types of faults (ag, ab, abc etc.) are identified. In the second case for a series compensated line a classifier is designed to identify whether a fault loop encounters the capacitor or not.

Application to a simple power network: A power system as shown in Fig.2 is considered for the fault classification objective. The voltage and current signals acquired at the relaying point are processed by the neural network to classify the faults in the transmission system. A conventional neural network approach to classify the fault type is preferred by a single network where all the three phase signals are used as input vector. Such an approach is shown in Fig.16. In the

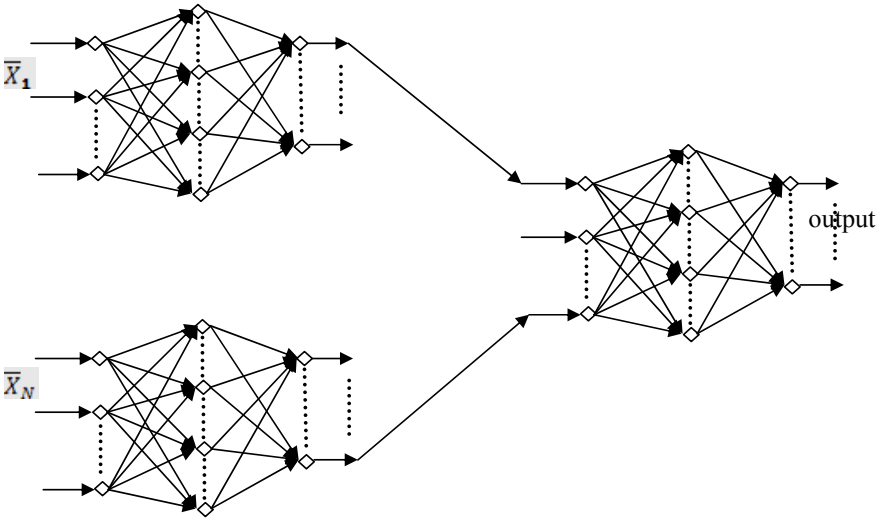


Fig. 15. A modular neural network structure

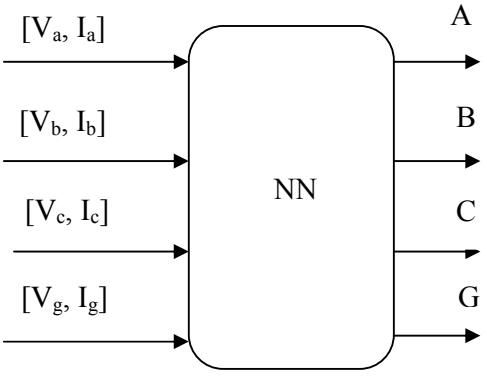


Fig. 16. Conventional neural fault classifier

modular structure, the networks are in parallel and their outputs represent the class of fault. The network output of a module should be 1 if the corresponding phase/ground is associated with ground else it should be 0. Each network takes voltage and current samples of that phase only. All the networks combine to provide the information on the type of fault and if no network output is 1 the classifier declares the state to be normal. This indicates that unlike conventional distance relaying scheme there is no requirement of separate fault detector in such classifiers. The structure of the classifier is shown in Fig.17 In this case NN-1 is for phase and NN-2 represents for ground.

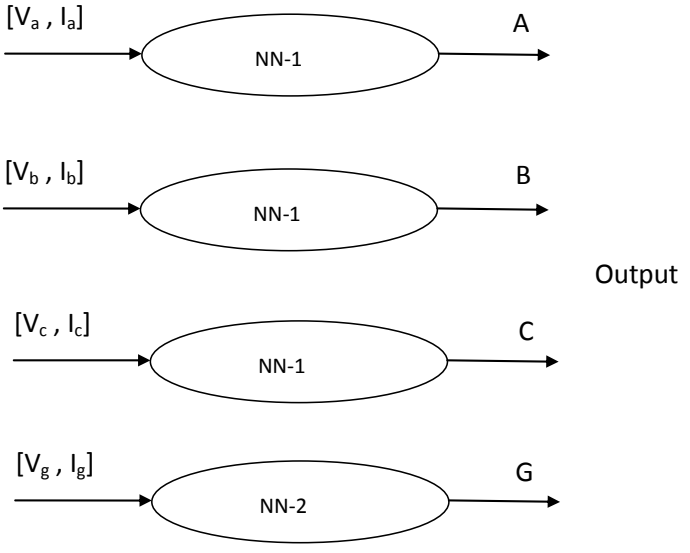


Fig. 17. The structure of the neural fault classifier

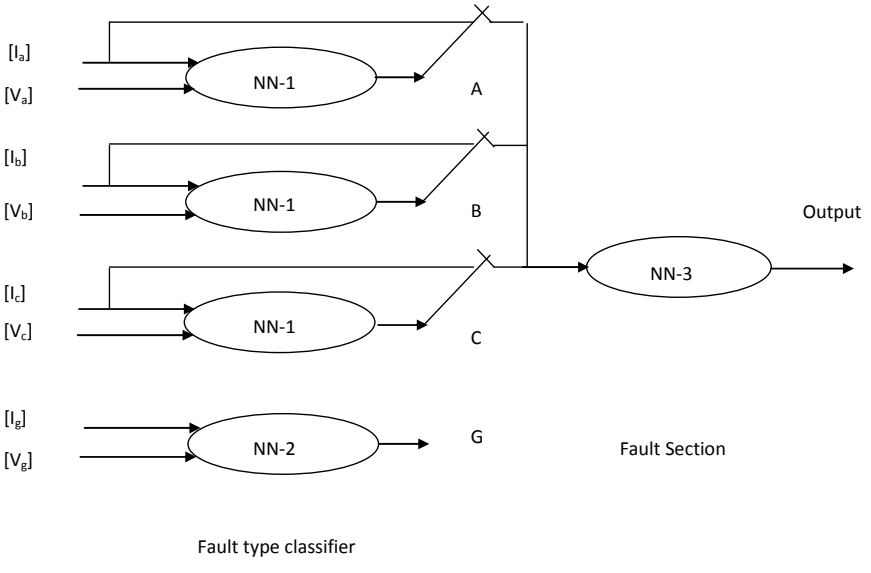


Fig. 18. The neural classifier for series compensated line

Application to a series compensated line: A series compensated network as shown in Fig. is next considered for classification challenge. In a series compensated line if the capacitor falls in the fault loop, as in the case of Fig.10, the fault distance estimation becomes difficult. There are two sections of the line connected through the series capacitor. In this case a different fault classifier is essential which discriminates the faults with regard to its originating section;

beyond the capacitor or within it. It is to be noted that fault type identification and section classification for a series compensated line also can be obtained by a single network in the conventional approach. With modular concept the simple solution to this classification problem is achieved by a structure as shown in Fig. 18. In this approach after fault type is identified a selector switch helps to provide data of any faulty phase current to the neural network NN-3 which classifies whether the fault is beyond the capacitor or within it. The classification by NN-3 is essential for a distance relay to derive the proper trip decision.

A typical result table for an AG fault in the line is provided in the table. The fault inception is at sample number 5 and NN-1 for phase-A and NN-2 provide involvement with the fault immediately after a sample which shows correct classification and consistency of the network outputs.

Table 3. Modular network performance

Sample no.	NN-1 outputs			NN-2 output (for ground)
	Phase-A	Phase-B	Phase-C	
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
→ 5	0	0	0	0
6	1	0	0	1
7	1	0	0	1
8	1	0	0	1
9	1	0	0	1
10	1	0	0	1

Applications to directional relaying

Directional relaying is widely applied for transmission line and other element protection. A directional relay is preferred when high speed relaying is necessary. It enhances the sensitivity and reliability of the protection schemes by discriminating the fault section with respect to the relay bus. The voltage and current waveforms, phasors or the derived sequence components are suitably used to estimate the fault direction. The fault currents lie in two distinct regions depending on the direction of fault which is shown in Fig. 19. There are number of directional relaying algorithms available using voltage and current waveforms and in many practical relay algorithms several logics are put together to derive the

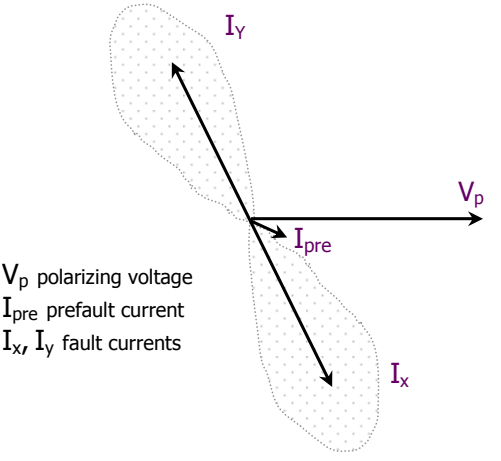


Fig. 19. Fault current position for different direction of faults

direction. However conventional algorithms find limitation due to the influence of prefault condition, close in fault, presence of decaying dc in the current signal and other signal processing issues. Considering fault direction estimation as a pattern classification problem, NN is suitably applied. It is demonstrated that the NN approach provides a way to adaptive protection and such a strategy is less affected by the presence of fault resistance, source impedance variation etc. A single network is used to identify the fault direction and the scheme is shown in Fig.20. The network is ANDed with the output of a fault detector for deciding the direction. Elman network with recurrent structure is also employed to obtain improved performance with time varying signals but with one structure.

In a three phase transmission system line, most of the faults are of line-to-ground fault type and the voltage and current signals of that phase are best representative for the direction estimation. By applying samples of three phase currents and voltages, the corresponding classification task becomes unnecessarily complex in most of the time. In conventional NN based approach direction identification is formulated as a 3-class problem; forward fault, reverse fault and

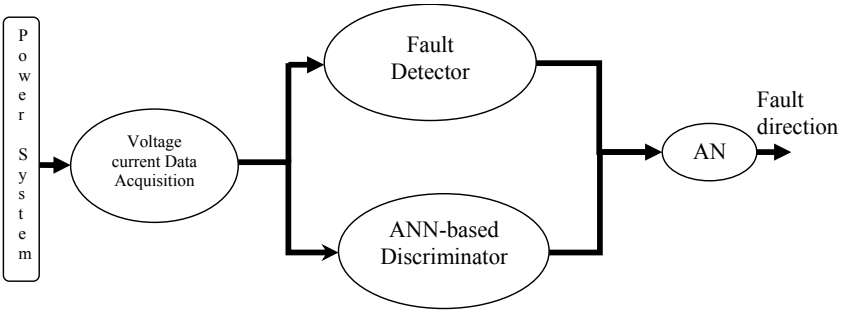


Fig. 20. Neural network N based directional relay

no fault. In this section a modular neural network based directional relaying is presented to obtain a reliable directional relaying algorithm which simplifies the complex task into number of subtasks.

The modular concept is introduced to directional relaying where the 3-class problem is divided into two 2-class problems. In this approach, a module discriminates forward fault only from other states and another module identifies reverse fault only, thereby the task is subdivided. Each module uses the corresponding phase current and voltage samples (for example, 5 samples each with a signal sampling rate of 1 kHz on a 50Hz system). This simplifies the neural network task further with redundant inputs being removed in the process. Further no fault detector is necessary. Thus the approach requires two modules of NN for each phase for the relaying algorithm.

The modular structure of NN for the directional relay is shown in Fig. 21. The relay consists of two modules of NN for each phase. Each NN1 module uses corresponding phase current and voltage samples to estimate whether the fault is in front of the relay or not. Similarly each NN2 module classifies the faults into two groups; fault behind the relay or not.

For each classifier module, several training patterns are generated by simulating normal state and faults in both directions. The fault situations simulated for such patterns are obtained by varying the fault location, inception angle, system frequency, initial load flow, and fault resistance. Each NN module is trained with an output value of '1' if the input vector corresponds to fault in its assigned direction else '0'. Such an approach reduces the size of the network including number of layers.

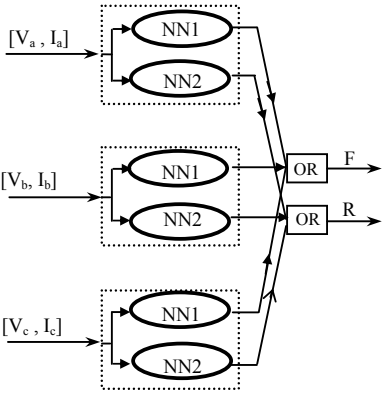


Fig. 21. Structure of the modular directional relay (F- forward, R- reverse)

6 Conclusions

Transmission line fault classification task is of more concern today as new compensating devices are being introduced and the systems are being operated close to their limits in the deregulated environment. This chapter highlights

applications of intelligent techniques such as neural network, fuzzy logic and neuro-fuzzy techniques for fault classifications of transmission lines. For series compensated network a wavelet based technique is also presented to identify the fault section using high frequency components of current signals. Each intelligent technique has its own merits and disadvantages. An integrated approach fuzzy-neuro technique is also provided in the text to exploit the advantages of both fuzzy logic and neural network. A modular concept in the neural network approach for fault classification in transmission systems is also presented. The modular concept simplifies a complex problem by dividing into smaller tasks. Each subtask is being accomplished by a neural network and then they being aggregated to provide the overall output. The inherent advantages modular structure includes higher accuracy, less training data and time and better structural interpretation. AI techniques are found to be powerful in classification of transmission line faults.

References

[1] Pradhan, A.K.: Adaptive Protection of Power Networks using Artificial Intelligent Techniques. Ph.D. Thesis, Sambalpur University, India (2001)
[2] Mohanty, S.R.: Detection and Classification of Transmission Line Faults. Ph.D. Thesis Indian Institute of Technology, Kharagpur, India (2006)
[3] Dash, P.K., Pradhan, A.K., Panda, G.: Application of minimal radial basis function neural network to distance protection. IEEE Trans. on Power Delivery 16(1), 68–74 (2001)
[4] Dalstein, T., Kulicke, B.: Neural network approach to fault classification for high speed protective relaying. IEEE Trans. on Power Delivery 10, 697–706 (1995)
[5] Song, Y.H., Xuan, Q.Y., Johns, A.T.: Comparison studies of five neural network based fault classifiers for complex transmission lines. Electric Power System Research 43, 125–132 (1997)
[6] Dash, P.K., Pradhan, A.K., Panda, G.: A novel fuzzy neural network based distance relaying scheme. IEEE Trans. on Power Delivery 15(3), 902–907 (2000)
[7] Wang, H., Keertipala, W.W.L.: Fuzzy-neuro approach to fault classification for transmission line protection. IEEE Trans. on Power Delivery 13(4), 1093–1103 (1998)
[8] Saha, M.M., Kastenny, B., Rosolowski, E., Izykowski, J.: First zone algorithm for protection of series compensated lines. IEEE Trans. on Power Delivery 16(2), 200–207 (2001)
[9] Youssef, O.A.S.: New algorithm to phase selection based on wavelet transform. IEEE Trans. on Power Delivery 17(4), 908–914 (2002)
[10] Chanda, D., Kishore, N.K., Sinha, A.K.: Application of wavelet multiresolustion analysis for identification and classification of faults on transmission lines. Electric Power Systems Research 73, 323–333 (2005)
[11] Pradhan, A.K., Routray, A., Pati, S., Pradhan, D.K.: Wavelet-fuzzy combined approach for fault classification of a series compensated transmission line. IEEE Trans. on Power Delivery 19(4), 1612–1618 (2004)
[12] Lu, B.L., Ito, M.: Task decomposition and module combination based on class relations: a modular neural network for pattern classification. IEEE Trans. Neural Networks 10, 1224–1255 (1999)

[13] Pradhan, A.K., Mohanty, S.R., Routray, A.: Neural fault classifier for transmission line protection - a modular approach. In: Power Engineering Society General Meeting, Montreal (2006)

[14] Sidhu, T.S., Singh, H., Sachdev, M.S.: Design, implementation and testing of an artificial neural network based fault direction discriminator for protecting transmission lines. IEEE Transactions on Power Delivery 10, 697–706 (1995); Sanaye Pasand, M., Malik, O.P.: High speed transmission system directional protection using an Elman network. IEEE Transactions on Power Delivery 13(4), 1040–1045 (1998)

[15] Lahiri, U., Pradhan, A.K., Mukhopadhyaya, S.: Modular neural network based directional relay for transmission line protection. IEEE Trans. on Power System 20(4), 2154–2155 (2005)

Fuzzy Reliability Evaluations in Electric Power Systems

Dusmanta Kumar Mohanta

1 Introduction

The fundamental operating feature of the Power system is that the electrical energy production and consumption are simultaneous. Therefore, the reliability requirement for power system is very high. The classical reliability evaluations of Power System adopts the widely accepted definition of reliability as the probability of a device performing its purpose adequately for he period of time indented under the operating conditions encountered [1]. In essence, for practical purpose, it is defined as the probability of a component serving in a given time period having a constant failure rate and expressed as exponentially distributed function. The classical reliability assessments are based on the probabilistic assumptions about state behavior and binary assumptions about state of the system, *i.e*. either in success state or in failed state. Hence classical reliability models are also known as probabilistic binary state (PROBIST) model [1-3].
In many cases, the classical probabilistic binary state (PROBIST) model representing binary state reliability indices such as mean time to failure (MTTF) and mean time to repair (MTTR) by expected (crisp) values seems to be inadequate from pragmatic prospective [4]-[6]. Most of the reliability data is obtained from databases. However, the data associated with equipment may not be exactly that exists in the database, either because it was not installed under the same conditions or just some new types of equipment are in vogue. Consequently, some uncertainties are associated with component indices due to the lack of up gradation of data. In addition, such failure and repair rates depend not only the component themselves, but also on other systematic factors that include company efficiency and operation policy. A fuzzy model takes into account both non-probabilistic as well probabilistic uncertainties related to mean time to failure/repair for the generating units incorporating the fuzzy MTTF and fuzzy MTTR based on probabilistic fuzzy state (PROFUST) model [4]-[5]. There has been a paradigm shift from probabilistic reliability assessments towards the possibilistic reliability assessments to take into account such uncertainties using probabilistic fuzzy state (POSFUST) model. Also the fuzzy Markov model uses the probabilistic approach with some enhancement using fuzzy concepts [4-9].

Dusmanta Kumar Mohanta
Professor, EEE Dept., Birla Institute of Technology
Mesra, Ranchi-835215, India
E-mail: dkmohanta@bitmesra.ac.in

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 103 – 130.
springerlink.com

The organization of this chapter is as follows. The section 2 depicts basic reliability concepts such as general reliability function, its exponential distribution, mean time to failure (MTTF), mean time to repair (MTTR), forced outage rat (FOR) etc. used in PROBIST model. Section 3 gives a brief overview of fuzzy reliability models such type-1 & type-II models in power system perspective. It depicts the type-I fuzzy reliability model (fuzzy load model) for hierarchical level (HC-1) of power system with case studies. Also it depicts the type-II fuzzy reliability models with case studies. Section 4 gives an overall conclusion of the chapter followed by references.

2 Basic Reliability Evaluation Concepts

For power system reliability evaluations the classical PROBIST model uses the general reliability function R(t). This expression is derived below in a more fundamental manner so as to have better appreciation of the underlying failure density function [1-3].

2.1 The General Reliability Function

Let us consider the case in which a fixed number N_o of identical components are tested, for which $N_s(t)$ = number surviving at time t and $N_f(t)$ = number failed at time t.

At any time t, the reliability R(t) is given by:

$$R(t) = \frac{N_s(t)}{N_o} = \frac{N_o - N_f(t)}{N_o} = 1 - \frac{N_f(t)}{N_o}$$

$$\text{Therefore, } \frac{dR(t)}{dt} = \frac{-1}{\text{No}} \frac{\text{d } N_f(t)}{\text{dt}}$$

As dt $\rightarrow 0$, $\frac{1}{\text{No}} . \frac{\text{d } N_f(t)}{\text{dt}}$ is defined as the instantaneous failure density f (t).

$$\text{Thus } \frac{\text{d R(t)}}{\text{dt}} = - \text{ f (t)}$$

The hazard rate $\lambda(t)$ is derived below to express reliability R (t) as a function of hazard rate.

$$\lambda(t) = \frac{\text{d N}_f(t)}{\text{dt}} \cdot \frac{1}{\text{N s (t)}} \quad \lambda(t) = \frac{\text{N}_o}{\text{N}_o} \frac{1}{\text{N}_s(t)} \cdot \frac{\text{d N}_f(t)}{\text{dt}} = \frac{\text{f (t)}}{\text{R (t)}} = \frac{- \text{ dR (t)}}{\text{R (t) dt}}$$

$$\text{Thus} \quad \int_o^t \lambda(t)\, dt = -\int_1^{R(t)} \frac{dr(t)}{R(t)}$$

$$\text{or,} \quad R(t) = e^{-\int_o^t \lambda(t)\, dt} \tag{2.1}$$

This equation gives the reliability as a function of time where the hazard or failure rate is also a function of time. No assumptions regarding the form of the failure function have been made in deriving this expression. If λ is a constant and therefore independent of time $R(t) = e^{-\lambda t}$.

Experience has shown that many components and particularly electronic components follow a relatively standard failure rate pattern with time. This basic pattern is shown in Figure 1 and is often referred to as the conventional bathtub curve.

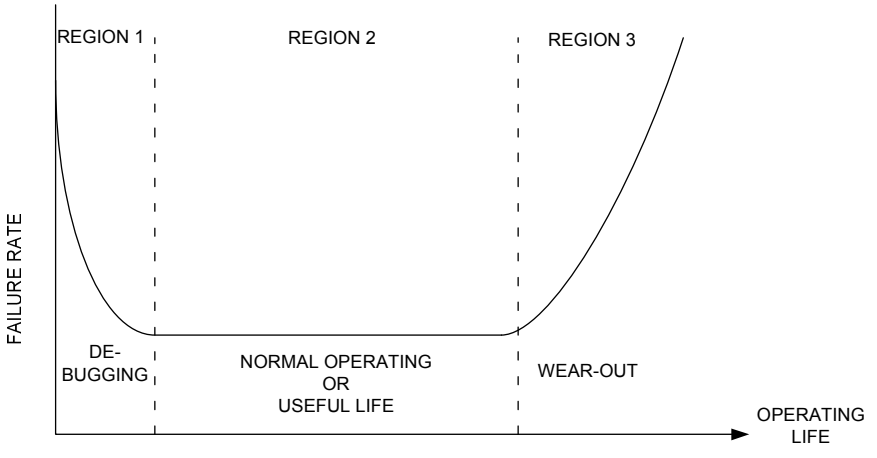


Fig. 1. Bathtub Curve

During the useful life or normal operating phase, the failure rate is constant and failures are assumed to occur purely by chance. The negative exponential distribution is valid in this region only. Region 1 in Figure1 is known as the infant mortality or de-bugging phase and could be due to manufacturing errors or improper design. The failure rate decreases as a function of time. Region 3 represents the wear-out or fatigue phase and is characterized by a rapidly increasing failure rate with time. The failure density in this region is often represented by the Normal Distribution. Other distributions such as the Gamma Distribution and the weibull distribution are also often in use [1-3].

Power system components such as generating units, transformers, switchgear, etc, can remain within the useful life period for the bulk of their economically feasible life by constant and careful preventive maintenance. In this way, insulation and mechanical elements are not allowed to enter an advanced wear-out state before they are replaced. This is an extremely important assumption, however, as reliability prediction based upon useful life rates is invalid and extremely optimistic if the system contains components which are operating within their wear-out period.

2.2 The Exponential Distribution

It has been shown previously that the probability of a component surviving a time t in a constant failure rate environment is given by R (t) = $e^{-\lambda t}$ and the failure density function

$$f(t) = \frac{-dR(t)}{dt}.$$ This is shown graphically in Figure 2.

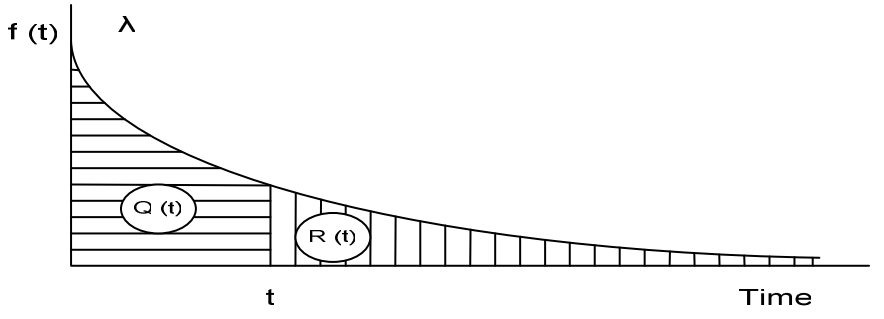


Fig. 2. Exponential Probability Density function

The probability of failure in time $t = Q(t)$

$$Q(t) = \int_{o}^{t} \lambda e^{-\lambda t} dt = 1 - e^{-\lambda t} \tag{2.2}$$

The probability of survival in time $t = R(t)$

$$R(t) = \int_{t}^{\infty} \lambda e^{-\lambda t} dt = e^{-\lambda t}$$

The expected value is often designated as the mean time to failure MTTF and in the useful life period is the reciprocal of the failure rate. A slightly different term Mean Time between Failure MTBF is used to indicate the cycle time between failures for systems that are repairable. The MTBF, therefore, exceeds the MTTF by a small margin attributable to the time associated with component repair. It is important to realize that the term Mean Time to Failure is normally applied only to

the useful life period. A component could be stated to have a mean wear-out life is very much less than the MTTF. This can be illustrated by using a numerical example. The mean wear-out life of a component is 1000 hours. These two statements are not contradictory. The average wear-out life of the component is 1000 hours. Assuming that the failure density function in the wear-out region is normally distributed, the time at which the component actually entered this region depends upon the standard deviation of this distribution. The larger the standard deviation, the shorter the useful life period. During the useful life period, failures occur by chance and the failure rate is constant. The failure density function is an exponential distribution with parameter λ, the failure rate. The reciprocal of this failure rate is the mean time to failure assuming that the distribution continues to be applicable. The MTTF can therefore by very much linger than the mean wear-out life.

2.3 Mean Time to Failure (MTTF)

The expected value of $E(x)$ for a continuous random variable with probability distribution function $f(x)$ is given by

$$E(x) = \int_0^\infty x\, f(x)\, dx$$

$$\text{So MTTF} = \int_0^\infty \lambda t\ e^{-\lambda t}\, dt$$

$$= 1/\lambda \tag{2.3}$$

$$\text{Similarly mean time to repair (MTTR)} = 1/\mu \tag{2.4}$$

Where μ is the repair rate.

2.4 Forced Outage Rate (F.O.R.)

The probability of finding the unit on outage at some future time is known as the forced outage rate (F.O.R.) where

$$F.O.R. = MTTR/(MTTR + MTTF) \tag{2.5}$$

3 Power System Fuzzy Reliability Models

Power system fuzzy reliability assessments can be broadly segregated into two categories, viz., type I and type II[4-8]. Type I fuzzy reliability calculations

are related to the uncertainty in defining the power consumption and therefore in the impacts of failures of supply. Fuzziness of a load value will be translated into uncertainty about the actual consequences of failures. The concept of fuzzy load is the edifice of such calculations, as explained in the subsequent section.

Unlike the classical PROBIST model is based on the probabilistic assumption about the system behavior along with the binary state assumption, the fuzzy reliability models are based on the fuzzy assumption about the system behavior along with either the binary state assumption or fuzzy state assumptions, leading to different fuzzy type II reliability models. The probabilistic fuzzy state (PROFUST) model assumes the probabilistic assumption state behavior, but introduces a fuzzy state assumption. Similarly the possibilistic fuzzy state (POSFUST) model combines the possibilistic assumption about state behavior with the fuzzy state assumption. On the other hand, the possibilistic binary state (POSBIST) model incorporates a possibilistic assumption about state behavior together with the binary state assumption. The following section corroborates such fuzzy reliability models with relevant case studies.

3.1 Type I Power System Fuzzy Reliability Evaluation Incorporating Load Forecasting Uncertainties

Power system reliability assessment has for long extended beyond continuity of supply. The impact on consumer supply has become an essential way of measuring the quality of system design and of comparing the merits of alternatives. The most classical measure of reliability is the loss of load probability (LOLP) and this reliability index in fuzzy paradigm is represented by fuzzy loss of load probability (FLOLP). There are two ideas that justify adopting fuzzy representations of loads in power system models for the purpose of reliability analysis. The first relates to the evaluation of fuzzy LOLP index of a system in order to understand how uncertainty in load prediction affects the reliability performance of a system. The second relates to the comparison between the alternatives.

Type I fuzzy calculations [4-9] at the hierarchical level 1(HL-1) are related to the uncertainty in defining the power consumption and therefore measuring the impacts of failure of supply. Fuzziness of load value is translated into uncertainty about the actual consequences of random failures of generating units. Although the relation between a changing load and the varying calculated index is far from linear, but still its characteristics are well captured by the fuzzy set description. In essence, the fuzzy representation of load in power system establishes how the uncertainty in load prediction affects the reliability indices of a system.

A fuzzy description of the load curve would be one that defines, at every level α, an interval of confidence for such curve. If, for instance, a crisp load

curve is given by $L = f(q)$, then $q = f-1(L)$. A fuzzy load curve could then be defined, at a level α, based on the crisp function f by equation (3.1), given below.

$$L\alpha = [(1-\Delta\alpha -) f(q); (1-\Delta\alpha +) f(q)] \quad (3.1)$$

with $\Delta\alpha -$ and $\Delta\alpha +$ being two non-strictly monotonically decreasing functions with α. Conversely, there will be a fuzzy description of the probability q through the equation (3.2).

$$q\alpha = [\max\{0, f-1(1/(1-\Delta\alpha -) L)\}; \min\{1, f-1(1/(1-\Delta\alpha +) L)\}] \quad (3.2)$$

where qα is the probability at a confidence level of α. For hierarchical level I (HL-I), the fuzzy loss of load probability (FLOLP) is given by equation (3.3).

$$FLOLP = \sum_{j} \tilde{p}[C-PD(j)]q[PD(j)] \quad (3.3)$$

where C = System effective installed capacity, taking maintenance of units into account; PD(j)= Load demand during jth period; $\tilde{P}[C-PD(j)]$ = fuzzy probability of capacity outage of [C-PD(j)] for a given period j; q = probability of load exceeding available capacity. For example, a captive power plant having six units, each of 120MW has an effective installed capacity of 720 MW when no unit under maintenance, but the effective installed capacity reduces to 600 MW when one unit is under maintenance.

3.1.1 Case Studies and Discussion of Results

Usually the captive power plants are synchronized with local state grid so that excess power can be exported to grid and at the time of exigencies, power can be imported as well. But during the grid failure, the captive power plants can be operated in the islanded condition till the grid revives its healthy state. This greatly enhances the reliability of the supply of the utility industries connected to captive power plants. For the purpose of stability after islanding, having a system frequency lower than the nominal frequency of 50 Hz, a minimum export or reserve Ro is maintained and this requirement is incorporated as a constraint in the problem formulation. For islanding operation there is a provision of an islanding circuit breaker (CB) as depicted in figure 8. In the schematic diagram an aluminum smelter is considered as the utility industry for a captive power plant consisting of 6 generating units each of 120MW for supplying uninterrupted power. Although the smelter is located nearby, there are two very short transmission lines (T.R. 1 & 2) as indicated in fig 11. The circuit breakers and bus bars are shown at appropriate locations for the case study of the proposed research work.

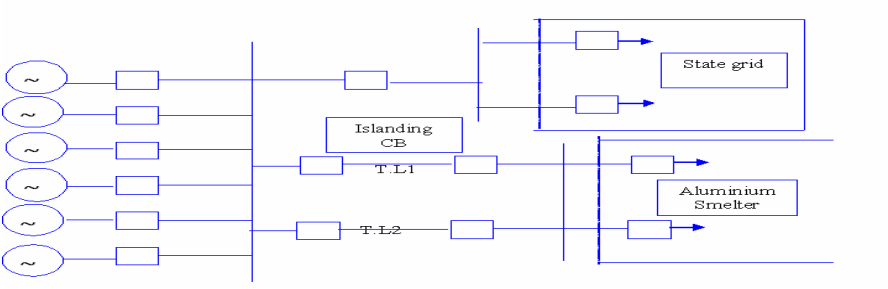


Fig. 3. Schematic diagram of a captive power plant

Legends:

: Generating Unit

: Circuit Breaker

: (T.L.1) Transmission line 1 & (T.L.2) Transmission line 2

: Feeders for State Grid & Aluminum Smelter

: Bus bar

The generating unit data pertaining to maintenance scheduling is given table 1.

Table 1. Optimum Maintenance Scheduling

Index of generating unit	Starting period of maintenance (Week)	Maintenance Class	Starting period of maintenance (Week)	Maintenance class	Starting period of maintenance (week)	Maintenance class
1	9	D	89	S	175	D
2	40	D	144	D	229	S
3	20	S	77	S	154	D
4	27	D	97	D	165	D
5	64	D	108	D	185	D
6	56	S	119	S	195	D

The type of maintenance is specified as S or D. The D class maintenance is for the detailed maintenance of boiler, turbine and generator. The duration of D class generator is for ten weeks. S class maintenance is a simplified maintenance for both boiler and generator, and it lasts for six weeks. History of the generating units MTTF and MTTR is given in Table 2.

Table 2. MTTF and MTTR of Generating Units

Index of Generating Unit	MTTF (hours)	MTTR (hours)
1	769	95
2	828	92
3	910	90
4	1156	87
5	1222	78
6	1560	65

Table 3 gives the load demand for the captive power plant for the whole planning horizon over 260mweeks.

3.1.2 Maintenance Scheduling: A Review

The goal of maintenance scheduling for thermal generating units of a captive power plant is to allocate a proper timetable for planned outage at a desired security level as required by a utility industry. The objective function (F) is formulated to minimize the sum of squares of reserve generation as given below.

$$Minimize\ F = \sum_{j=1}^{J} \{ \sum_{i=1}^{\max unit} P\ [i]\ [j] - PM\ [i]\ [j] - PD\ [j]\ - Ro \} 2 \tag{3.4}$$

Where P[i][j] is the generating capacity of ith unit in the period j, PM [i][j] is the power loss due to maintenance of unit I during period j and R0 is the minimum reserve requirement.

The following constraints are considered for the captive power plant maintenance-scheduling problem.

Starting Period of Maintenance

Keeping statutory safety regulations in view, the thermal units must be maintenance within a maximum permissible period Qmax [i] [nm]. The earliest staring period Qmin [i] [nm] is chosen from the general consideration that the thermal unit maintenance must commence within one year plus or minus one month from last maintenance. So in equation (3.5) below, term 48 weeks has been added.

$$Qmin\ [\ i]\ [nm] = Q\ [i]\ [nm\text{-}1] + M\ [(S[i][nm\text{-}1])] + 48 \tag{3.5}$$

Usually the maximum permissible limits are governed by following statutory norms as follows. Unit that has not been maintained over 1.5 years (Turbine over 156 weeks)

Table 3. Load demand PD [j] for a planning horizon of 20 weeks

j (week)	PD[j] (MW)	j (week)	PD[j] (MW)	j (week)	PD[j] (MW)	j (week)	PD[j] (MW)	j (week)	PD[j] (MW)	j (week)	PD[j] (MW)
1	412	45	418	89	437	133	435	177	426	221	416
2	406	46	435	90	413	134	409	178	421	222	425
3	420	47	437	91	416	135	432	179	408	223	433
4	411	48	418	92	439	136	416	180	426	224	422
5	413	49	409	93	407	137	437	181	424	225	411
6	432	50	439	94	440	138	408	182	411	226	425
7	444	51	430	95	424	139	441	183	438	227	436
8	411	52	432	96	414	140	407	184	420	228	415
9	445	53	432	97	449	141	403	185	432	229	411
10	447	54	400	98	424	142	432	186	449	230	422
11	432	55	432	99	429	143	432	187	441	231	433
12	414	56	424	100	426	144	403	188	420	232	444
13	431	57	433	101	419	145	405	189	421	233	425
14	438	58	446	102	415	146	443	190	443	234	422
15	435	59	438	103	424	147	405	191	416	235	412
16	405	60	433	104	413	148	427	192	434	236	413
17	450	61	441	105	425	149	449	193	446	237	412
18	442	62	429	106	432	150	405	194	438	238	422
19	407	63	441	107	431	151	442	195	438	239	422
20	410	64	412	108	447	152	448	196	433	240	433
21	403	65	421	109	424	153	443	197	442	241	415
22	433	66	407	110	444	154	433	198	403	242	412
23	415	67	432	111	407	155	406	199	415	243	426
24	429	68	410	112	428	156	422	200	422	244	433
25	432	69	426	113	420	157	415	201	433	245	415
26	412	70	417	114	437	158	409	202	409	246	416
27	411	71	429	115	429	159	450	203	416	247	425
28	432	72	442	116	405	160	430	204	444	248	422
29	425	73	438	117	418	161	429	205	428	249	444
30	406	74	427	118	447	162	449	206	435	250	433
31	409	75	426	119	413	163	410	207	449	251	411
32	406	76	412	120	408	164	436	208	420	252	415
33	439	77	422	121	421	165	412	209	415	253	418
34	417	78	438	122	407	166	420	210	425	254	422
35	407	79	441	123	412	167	400	211	433	255	417
36	427	80	408	124	414	168	422	212	444	256	411
37	419	81	440	125	404	169	441	213	425	257	408
38	425	82	412	126	426	170	412	214	440	258	421
39	442	83	431	127	412	171	422	215	412	259	428
40	401	84	409	128	425	172	430	216	432	260	415
41	433	85	416	129	437	173	444	217	421		
42	446	86	448	130	417	174	413	218	412		
43	432	87	435	131	400	175	404	219	413		
44	416	88	446	132	433	176	430	220	415		

must be maintained through a detailed maintenance and the next maintenance must be completed within 1.5 years. Simplified maintenance for boiler and generator must be completed within 1.5 years (turbine within 156 weeks). No extensions are permitted before or after minor maintenance. The maintenance class adopted before or after the minor maintenance must be a detailed one. The minor maintenance cannot be adopted continuously for the same unit. The minor maintenance is carried out so that more time-consuming maintenance can be postponed. These statutory norms are given for calculation of maximum permissible extension F[i][nm] . Q[i][nm] is calculated according to the equation given below.

$$Qmax\ [\ i]\ [nm] = Qmin\ [i]\ [nm] + F\ [i]\ [nm] \quad (3.5.1)$$

$$\text{All the time reserve requirement } R\ [j]\ \geq Ro \quad (3.5.2)$$

where R_0 is the minimum reserve requirement and taken to be 80 MW.

A captive power plant for an aluminum smelter, represented by Figure 3.8 in section 3.5, is chosen for case studies having 6 thermal units, each of 120 MW. The maintenance scheduling for optimization of safety and reliability has been formulated based on deterministic levelized reserve method, as described in Chapter III. The statutory safety norms in terms of extension terms F [i][nm] have been specified in Table 1. The data pertaining to generating units has been given in Table 2. The load demand PD [j] for each period j has been given for the whole planning horizon spanning over 260 weeks in Table 3.6. The optimum results obtained for levelized reserve method using hybrid GA/SA is given as Table 5.

The fuzzy load model is basically due to lack of precision in defining the load and given as Figure.

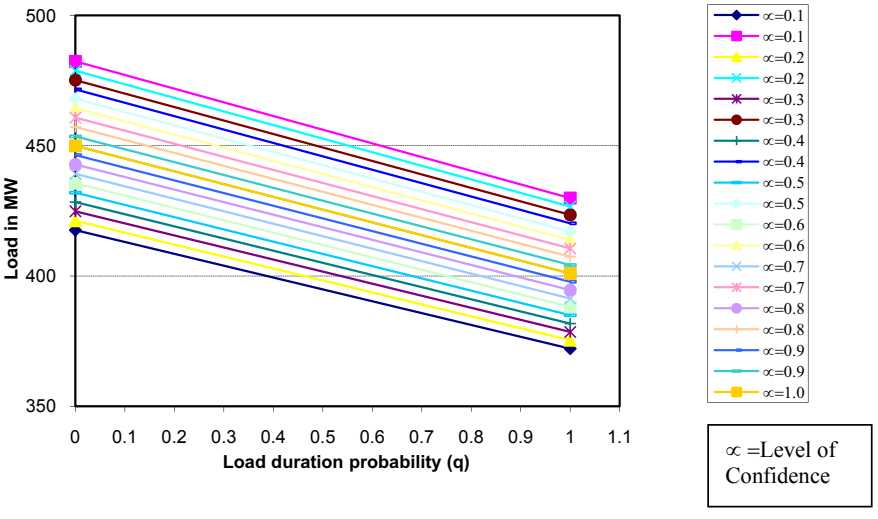


Fig. 4. Fuzzy load curve

Since the load variation is between 401 MW to 450 MW, an approximate monotonically decreasing crisp cumulative load duration curve is described by L=450-49q, where q is the probability of load exceeding available capacity. The fuzzy load curve is defined in an interval L ± 8% and the fuzzy load at any interval of confidence α, namely L∝, is given by the following equation based on equation (5.1). For the case study, the different values of L∝ are found out using equation (3.4).

$$L\propto = [(1-(0.08-0.08\propto))(450 - 49q) ; (1+(0.08-0.08\propto))(450 - 49q)] \quad \ldots\ldots \quad (3.4)$$

The fuzzy load is computed using equation (3.4) for different values of confidence level α, by considering 8% uncertainty in defining the load duration curve and the fuzzy load curve as given in Figure 4. For fuzzy load represented by L – 8%, the $L\propto$ has two values corresponding to q=0 and q=1 respectively. For example, at $\propto$ = 0.1, the respective values are 417.600006 MW and 372.127991MW. Similarly for fuzzy load represented by L + 8%, the $L\propto$ has two values corresponding to q=0 and q=1 respectively. For example, at $\propto$ = 0.1, the respective values are 482.399994 MW and 429.872009MW.

The fuzzified load for different confidence levels is given in Table 5.1. The contribution of outage stage $q\propto$ is calculated using equation (3.2). This fuzzy value $q\propto$ is multiplied with the corresponding capacity outage probability, obtained from the capacity outage Tables 3-4, to calculate the fuzzy loss of load probability (FLOLP) index based on equation (3.3).

The FLOLP values computed for the two cases, namely, when no unit is under maintenance as well as when one unit is under maintenance are given in Figure 5 and Figure 6 respectively.

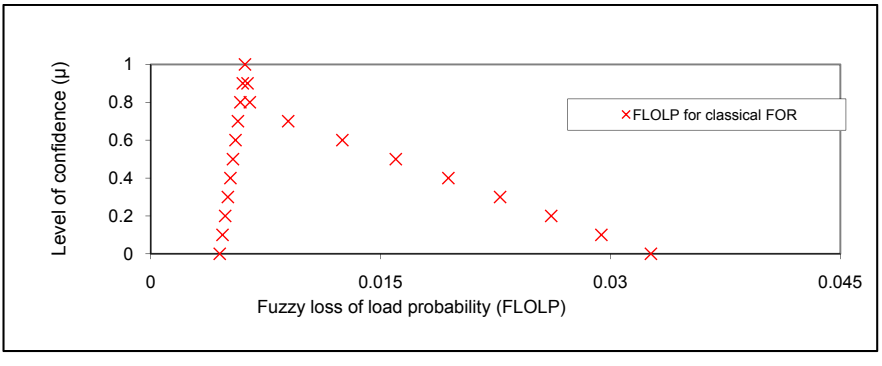


Fig. 5. Fuzzy loss of load (FLOLP) when no unit is under maintenance

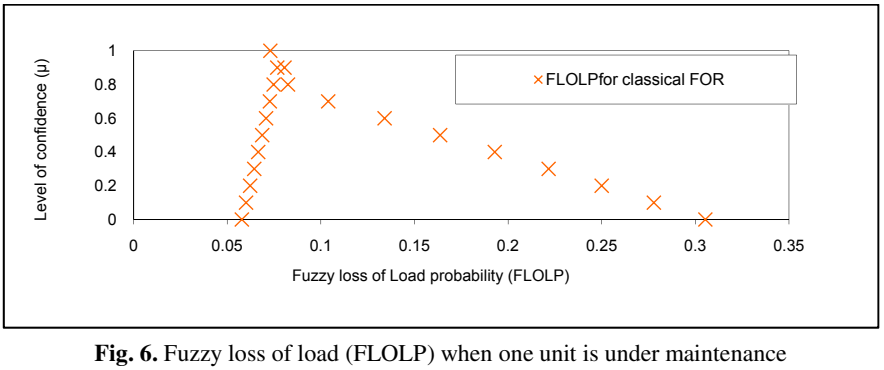


Fig. 6. Fuzzy loss of load (FLOLP) when one unit is under maintenance

When no unit is under maintenance and the effect of uncertainties related to load forecasting are considered, the FLOLP has a range of 0.04501 to 0.032639 for the maximum load uncertainty corresponding to level of confidence α having value of zero, as shown in Figure 5. Similarly for α=1, the FLOP has a crisp value of 0.00616027 when the uncertainties related to load forecast are considered in fuzzy model. When one unit is under maintenance, the uncertainties due to load forecasting yield FLOLP having a range of .057766 to 0.305531 when α= 0 and 0.073 when α= 1, as given in Figure 6. Thus the FLOLP range denotes the risk limits associated with the uncertainties.

The load forecasting uncertainties incorporated through the fuzzy model corroborate the fact that a small uncertainty, represented by a linear variation of the load predicted, has an ample effect on the reliability index uncertainties. Also it is evident that although the relation between predicted load and a varying FLOLP is not linear, but the model gives some boundary of variation. Since captive power plants cater to power sensitive utilities, the boundary of variation of reliability indices is very helpful for user industries to assess risks when the load requirements are not properly defined. But in many cases, the notions of component reliability indices such as failure rate or mean repair time by crisp numbers are not adequate from pragmatic prospective and therefore the present research uses different fuzzy models under type II reliability assessments in subsequent sections.

3.2 Type 2 Power System Fuzzy Reliability Models

The fuzzy reliability models are based on the fuzzy assumption about the system behavior along with either the binary state assumption or fuzzy state assumptions, leading to different fuzzy type II reliability models. The probabilistic fuzzy state (PROFUST) model assumes the probabilistic assumption state behavior, but introduces a fuzzy state assumption. Similarly the possibilistic fuzzy state (POSFUST) model combines the possibilistic assumption about state behavior with the fuzzy state assumption.

3.3 Probabilistic Fuzzy State (Profust) Model for Incorporation of Uncertainties in Forced Outage Rate of Generating Units

The classical reliability assessments are based on the probabilistic binary state model. Such a model presumes probabilistic assumptions about state behavior and binary assumptions about state of the system. The Reliability Test Systems (RTS) published by the IEEE Application of Probability Methods (APM) subcommittee has been a valuable and consistent reference source [78] in reliability assessment of generation systems and composite systems for utilities, consultants and universities. But in many cases, the classical probabilistic binary state (PROBIST) model [62] representing binary state reliability indices such as forced outage rate

(FOR) by crisp number seems to be inadequate from pragmatic perspective. Most of the reliability data are obtained from databases. But the data associated with equipment may not be exactly that exists in the database; either because it was not installed under the same conditions or just some new types of equipment is in vogue. Consequently some uncertainties are associated with component indices due to lack of upgradation of data. A probabilistic fuzzy state (PROFUST) model [62] takes into the account the uncertainties related to forced outage rate by considering the uncertainties related to mean time to failure (MTTF) and mean time to repair (MTTR) for the generating units. There has been a paradigm shift from probabilistic reliability assessments towards the possibilistic reliability assessments to take into account such uncertainties in the Risk Analysis paradigm [60,61], but type II PROFUST model uses the probabilistic approach about state behavior, but fuzzy state of generating units using fuzzy concepts.

The hybrid fuzzy model seems to be quite promising in reliability analysis of captive power plant maintenance scheduling problem. This is a long-term scheduling problem of planned outages for regular maintenance [10,16]. For such a long term operational planning the uncertainties in the planning variables such as forced outage rate (FOR) of generating units have to be considered in order to produce a realistic plan [9-16]. In system reliability evaluation it is frequently assumed that the failure and repair rates are constant during usual system operation without actually substantiating the assumptions. For a complex problem like operation and maintenance of a power station generating unit, constant failure and repair rates are inadequate models. Since diagnosis and maintenance tend to be experience-based skills, the expert opinion [6-9] is very crucial to assign fuzzy values to mean time to failure (MTTF) and mean time to repair (MTTR) for generating units in the fuzzy hybrid model to obtain fuzzy FOR. Such fuzzy-FOR reflects the condition of operation and maintenance of thermal units more realistically for long-term operation planning like maintenance scheduling.

This section of dissertation corroborates the influences of the uncertainties of forced outage rate of generating units as well as that due to load on optimum captive power plant maintenance schedules obtained through intelligent computational techniques as discussed in Chapter IV.

The mean-time-to failure (MTTF) of a generating unit usually decreases with increase of age of the unit. Also, it depends on the type of previous maintenance and the MTTR increases as the time gap increases with respect to previous maintenance. But there is inherent uncertainty in defining MTTF and MTTR. In order to represent uncertainty in expert system, the analysis begins from a fundamental model of uncertainty based fuzzy mathematics [6-9] and leads to a rule-based expert system development for effectively extracting information from available data that leads to coherent conclusion. The knowledge of the experts is represented by a large number of rules so that the overall uncertainty must be calculated to reach conclusion. As in any rule-based system, the rules are chained together by what is called the inference engine. The important consideration of the inference engine is the methods by which the uncertainties are propagated among rules in the reasoning process.

The predominant factors leading to uncertainties regarding MTTF as well as MTTR are the age of the generating units, type of previous maintenance and time gap from previous maintenance. Usually the MTTF decreases with aging of units. Also an increase of duration from previous maintenance causes a decrease of MTTF. But the MTTF increases after maintenance and the effect is more prominent in case of class D maintenance compared to class S maintenance. The reverse is the effect in case of MTTR. It increases with the increase of the age as well as gap from previous maintenance. The MTTR decreases after maintenance and the effect is more prominent in case of class D maintenance compared to class S maintenance. The factors affecting MTTR and MTTF are taken as input to the fuzzy model. The universe of discourse for each fuzzy variable is quantified into a number of overlapping fuzzy sets, known as linguistic variables. For the input variable 'age of the generating units', the universe of discourse is taken as 20 years and is quantized into linguistic variables A1 to A20 for MTTF and A1 to A10 for MTTR. Similarly for the second input, 'number of weeks from previous maintenance', defined over a universe of discourse of 105 weeks is quantized into 10 linguistic variables, namely from G1 to G10 for both MTTF and MTTR. The third input variable 'type of previous maintenance' is quantized into 8 linguistic variables, namely, D1 to D4 for D class maintenance and S1 to S5 for S class maintenance. The output for MTTR is quantized into 8 linguistic variables, namely MTTR1 to MTTR8 and output for MTTF from MTTF1 to MTTF18. The inputs are combined through the rules given by experts to give fuzzified output from the range of values from MTTF1 to MTTF18. For example, when the age corresponds to A1, gap to G1 and maintenance type to D1, then output is MTTF2. Based on expert opinion, all possible rules are framed and then the output is defuzzified using centriod method to get the crisp values of expert FOR. Similarly the mean time to repair (MTTR) for a unit increases with aging of units and also with progression of gap with respect to the previous maintenance. Also the type of previous maintenance affects MTTR. Suitable triangular membership functions are ascribed to the input and output variables in fuzzy model to take the uncertainties into account and rules have been chained together based on expert opinion to obtain MTTR and MTTF. The output fuzzy values are defuzzified using centriod method. Based on the values of MTTR and MTTF, the expert fuzzy FOR is found using the equation (3.5).

$$\text{Expert fuzzy forced outage rate (fz-FOR)} = \text{MTTR}/(\text{MTTR}+\text{MTTF}) \qquad (3.5)$$

3.3.1 Case Study with Numerical Results

A captive power plant for an aluminum smelter, represented by Figure ----- is chosen for case studies having 6 thermal units, each of 120 MW and the optimum maintenance schedule obtained from levelized reserve method using hybrid GA/SA is given in Table 4.5. The effect of uncertainties related to forced outage rate (FOR) of generating units on the optimum schedule is investigated using the PROFUST model and the associated risk even with an optimum schedule is quantified as a fuzzy index, known as fuzzy loss of load probability (FLOLP).

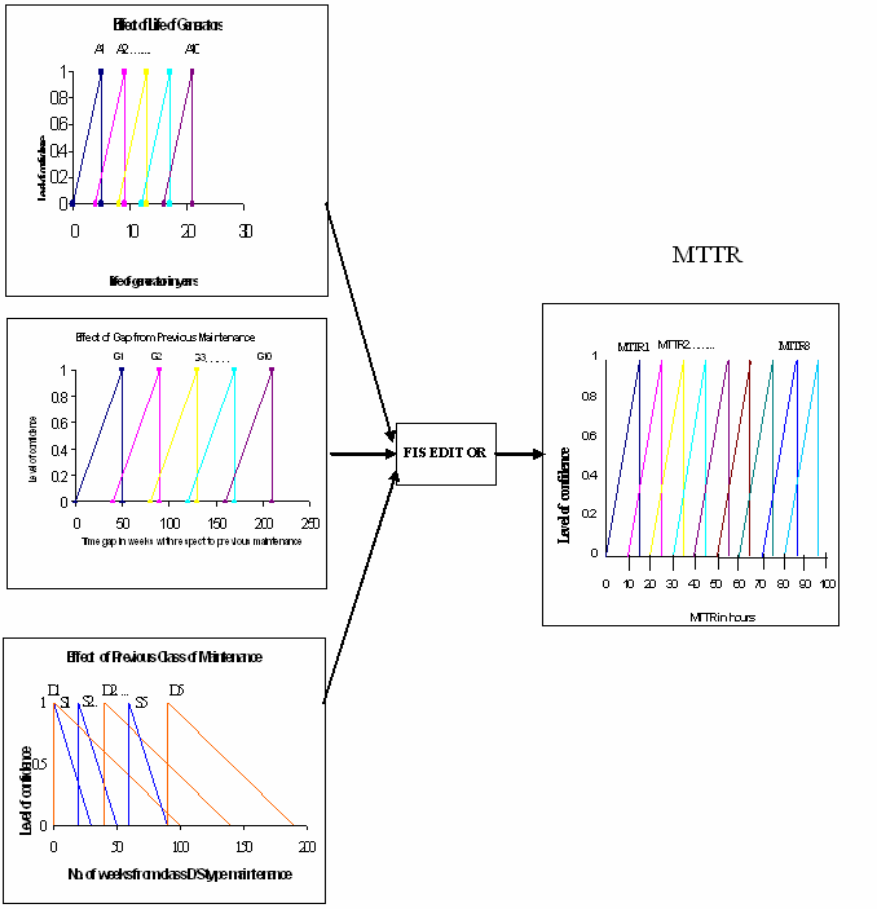


Fig. 7. A schematic diagram for fuzzy model for MTTR

Since the gestation period for the thermal power plants cannot be ignored, therefore although the units are similar, they differ with respect to age so that the MTTF and MTTR are also different for each generating unit. The data for the age of the units given in Table 2 is used in PROFUST model for finding out MTTR and MTTF to calculate fuzzy FOR. The values of MTTR and MTTF are found out just before maintenance as well immediately after maintenance. For example, 'the periods preceding maintenance count nm=4' represents the values of MTTR and MTTF just before maintenance and similarly 'the periods following maintenance count nm=4' represents the values immediately after maintenance is carried out.

A schematic diagram for probabilistic fuzzy state model for MTTR is given in Figure 7. The MTTR for a unit increases with aging of units and with progression of gap with respect to the previous maintenance. The MTTR decreases after maintenance and the effect is more prominent in case of D class maintenance compared to S class maintenance. The triangular membership functions with

increasing trend are used for two fuzzy input variables 'age of the generating units' and 'number of weeks from previous maintenance'.They are quantized from A1 to A10 and from G1 to G10 respectively. The third input variable 'type of previous maintenance' is represented by triangular membership with decreasing trend and quantized into D1 to D4 for D class maintenance and S1 to S5 for S class maintenance. The output variable MTTR is represented by triangular membership with decreasing trend and quantized into MTTR1 to MTTR8. Based on expert opinion, rules are framed for the FIS editor (Mamdani type) and the output MTTR is obtained. The defuzzification of output is done to get the crisp values of MTTR. The MMTF values are obtained for the whole planning horizon in the similar manner based on expert opinion.

Usually six values of FOR are found in the planning horizon that precede and follow the 3 scheduled maintenances, i.e., nm=3, 4 and 5. Although ideally there should have been 36 case studies corresponding to 18 cases when no unit is under maintenance and 18 cases when one unit is under maintenance [20], but the FOR values do not differ much for an individual unit compared to change of FOR from one unit to another when their effects on outage probabilities are considered. Therefore the best estimated value of expert FOR is considered for each unit.

The outage table for the case having no unit under maintenance is given as Table 3 and Table 4 corresponding to cases when no unit is under maintenance and one unit is under maintenance respectively, based on classical model. The outage table based on expert fuzzy FOR is given as Table 5.4.

The load forecasting uncertainties are incorporated through the fuzzy load model, as corroborated in section 5.1 and fuzzy load curve depicted as Figure 4 in section 5.1.1. The fuzzified load for different confidence levels is given in Table 5.1.The contribution of outage stage $q_{\propto}$ is calculated using equation (3.2). This fuzzy value $q_{\propto}$ is multiplied with the corresponding outage stage probability based on expert FOR, from Table 5.4, to calculate the fuzzy loss of load probability (FLOLP) using equation (3.3). The FLOLP so obtained for the optimum schedule obtained using hybrid GA/SA technique based on classical constant FOR values are compared with that obtained using expert fuzzy FOR values. For both the cases, uncertainties related to load forecasting are taken into account. The FLOLP computed for the two cases, viz., when no unit is under maintenance as well as when one unit is under maintenance is given in Figure 8 and Figure 9 respectively.

When no unit is under maintenance and the effect of uncertainties related to FOR are considered, the FLOLP has a range of 0.030761 to 0.131796 for the maximum load uncertainty corresponding to level of confidence α having value of zero, as shown in Figure 8. But when the uncertainties related to FOR of generating units are ignored, for the same value of α, the FLOLP has a range of 0.04501 to 0.032639 indicating lesser range of LOLP as uncertainties are reduced. Similarly, for α=1, the FLOLP has a crisp value of 0.0415125 when the uncertainties related to FOR values are considered in fuzzy model. For the same case, the crisp value becomes 0.00616027 when the uncertainties related to FOR

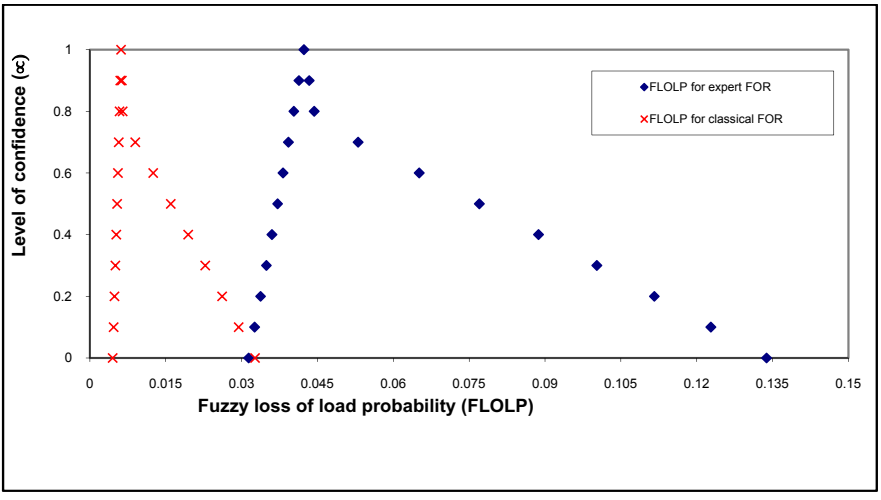


Fig. 8. Comparison of FLOLP based on expert and classical values of FOR when no unit under maintenance

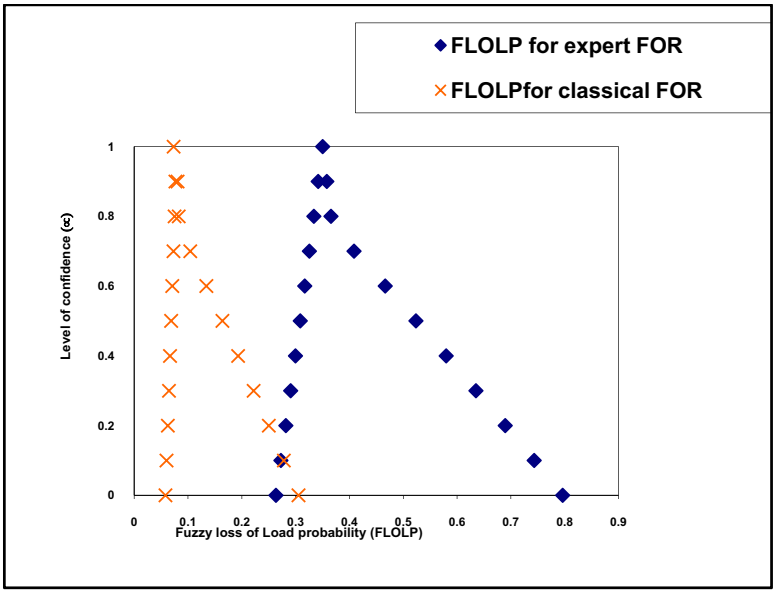


Fig. 9. Comparison of FLOLP based on expert and classical values of FOR when one unit under maintenance

are ignored in the classical probabilistic model. Thus, the FLOLP range denotes the risk limits associated with the uncertainties and it is seen that more is the uncertainty, greater is the range of FLOLP due to greater fuzziness. A similar trend is found when one unit is under maintenance, as given in Figure 9.

The PROFUS T model seems to be promising compared to classical PROBIST model for reliability analysis based on system adequacy for captive power plant maintenance scheduling as the uncertainties related to FOR are well captured in fuzzy model incorporating expert opinions. The additional impact of the uncertainties in the load represented by fuzzy load model is also incorporated in hybrid fuzzy model to calculate FLOLP. The reliability index FLOLP reflects uncertainties in terms of quantified risk and this is of immense use for planning related to captive power plants because the degree of acceptable risk for a captive power plant is implicitly related to the requirements of the process industry that is serving as the load for the captive plant. Also the advantage of using the probabilistic fuzzy model is that it is not exactly the replacement of the old faithful classical probabilistic model but an enhancement for better engineering approximation of FOR of generating units reflecting the impact of aging of units as well as the type of maintenance for getting a realistic schedule.

3.4 Possibilistic Fuzzy State (Posfust) Model for Incorporation of Uncertainties in Forced Outage Rate Of Generating Units

The standard approaches of reliability engineering rely on the probability model, which is often inappropriate for this task [1-3]. Probability based analyses usually require more information about the system than is known. Commonly, this results in inappropriate assumptions about the original data. Thus, any single value or distribution applied to the failure characteristics is likely to give a result that is misleading. Fuzzy logic offers an alternative to the probability paradigm using possibility paradigm[10-16]. The possibilistic approach for the system behavior allows for quantitative reliability evaluations that preserve the uncertainty in original data. The possibility model deals with uncertainty in a way that avoids making unwarranted conclusions and makes the consequences of the required assumptions more clear. The possibilistic approach based on possibility mathematics is a special branch of evidence theory and therefore a brief exposition of related topics such as fuzzy measures, evidence theory etc. are presented in the subsequent sections.

3.4.1 Fuzzy Measures

A fuzzy measure describes the vagueness or imprecision in assignment of an element to two or more crisp sets. Such a measure tries to describe the vagueness or imprecision in assigning an element to any of the crisp sets on the power set. This notion is not about the randomness because the crisp sets have no uncertainty about them. But the uncertainty is about the assignment. This uncertainty is usually associated with evidence to establish an assignment. The evidence can be completely lacking for total ignorance or the evidence can be complete for a probability assignment. Hence, the difference between a fuzzy measure and a

fuzzy set on a universe of elements is that in the former case the imprecision is in the assignment of an element to one of two or more crisp sets while for the latter the imprecision is in the prescription of the boundaries of a set.

3.4.2 Fuzzy Sets and Possibility Theory

Possibility theory can be formulated in terms of fuzzy sets. This alternative formulation of possibility theory is suggestive since fuzzy sets, similar to possibility bodies of evidence, are also based on families of nested sets, the appropriate α-cuts.

Possibility measures are directly connected with fuzzy sets via the associated possibility distribution functions. To explain this connection, let $\mathcal{V}$ denote a variable that takes values in a universal set V, and let the equation $\mathcal{V} = v$, where v ∈ V, be used for describing the fact that the value of $\mathcal{V}$ is v.

A fuzzy set F may be considered on V that expresses an elastic constraint on values that may be assigned to $\mathcal{V}$. Thus given a particular value v∈ $\mathcal{V}$, F (v) is interpreted as the degree of compatibility of v with the concept described by F. On the other hand, given the proposition "$\mathcal{V}$ is F" based upon F, it is more meaningful to interpret F (v) as the degree of possibility that $\mathcal{V} = v$. That is, given a fuzzy set F on V and the proposition "$\mathcal{V}$ is F," then possibility, rF (v), of $\mathcal{V} = v$ for each v ∈ V is numerically equal to the degree F (v) to which v belongs to F as given by equation (3.6).

$$rF(v) = F(v) \tag{3.6}$$

for all v ∈ V.

Function rF : V → [0, 1] defined by (5.23) is clearly a possibility distribution function on V. Given rF, the associated possibility measure, PosF, is defined for all A ∈ P (V) by equation (3.7).

$$\mathrm{Pos}_F(A) = \sup_{v \in A} r_F(v) \tag{3.7}$$

This measure expresses the uncertainty regarding the actual value of variable $\mathcal{V}$ under incomplete information given in terms of the proposition "$\mathcal{V}$ is F." For normal fuzzy sets, the associated necessity measure, NecF, is calculated for all A ∈ P (V) by equation (3.8).

$$\mathrm{NecF}(A) = 1 - \mathrm{PosF}(\bar{A}) \tag{3.8}$$

As an example, let variable $\mathcal{V}$ be temperature measured in oC and assume that only its integer values are recognized (i.e., V = Z). Let information about the actual value of $\mathcal{V}$ be given in terms of the proposition "$\mathcal{V}$ is around 21oC" in which the concept around 21oC is expressed by the fuzzy set F. This incomplete

information induces a possibility distribution function rF that, according to (3.6), is numerically identical with the membership function F. The α-cuts of F, which are nested, play the same role as the focal elements in possibilistic bodies of evidence formulated within evidence theory. Thus focal element and α-cuts correspond to each other in the two formulations of possibility theory. In the given example, the α-cuts (or focal elements) are A1 = {21}, A2 = {20, 21, 22}, A3 = {19, 20, 21, 22, 23}. Using equation (3.7) it is readily found that Pos (A1) = Pos (A2) = Pos (A3) = 1 and Pos (A1) = 2/3, Pos (A2) = 1/3, Pos (A3) = 0. Then, using equation (3.8), the results obtained are Nec (A1) = 1/3, Nec (A2) = 2/3and Nec (A3) = 1.

3.5 Possibilistic Approach for Evaluation of Fuzzy Forced Outage Rate (FOR)

The uncertainties related to forced outage rates (FOR) of generating units are important for studies related to long term generation cum expansion planning. But probability based analyses usually require more information about the system for parameters such as mean failure rates, failure distributions etc. Commonly, this sort of lack of information results in less appropriate assumptions about the original data. Thus, any single value or distribution applied to the failure characteristics is likely to give a result that is misleading. Fuzzy logic offers an alternative to the probability paradigm by using possibility paradigm. The possibilistic approach for the system behavior allows for quantitative reliability evaluations that preserve the uncertainty in original data, as explained in the previous section. In addition, experts in the field understand these uncertainties better. Therefore, the expert evaluation is modeled with fuzzy set theory based on possibilistic approach, as depicted below.

Notations:

t	time (a discrete variable) in weeks
T	fuzzy set whose members are various values of t
μT(t)	membership function of T
rf (t) , rr(t)	possibility distribution function {unit [fails , is repaired] at t}, t Є [0, ∞]
∏f (t) , ∏r (t)	cumulative possibility distribution {unit [fails , is repaired] in [0, t]}
Kf (t) , Kr (t)	cumulative possibility distribution {unit is [operating, not repaired] in [0,t]}

Assumptions:

1. Time t Є N (natural numbers). A similar analysis can be done for t Є R+ (positive real numbers).
2. A generating unit in the example has two states.

Let there be a fuzzy number T with μT(t) Є [0,1] , t Є [0 ,∞] representing an expert's opinion on the breakdown of generating unit at various values of t . The membership function represents rf (t) of a generating unit based on expert opinion, or in simple terms, the subjective expert opinion on possibility distribution function, e.g., breakdown of a unit at time t.

$$rf(t) = \mu T(t) \tag{3.9}$$

Such a definition of the relationship between the membership grade of a fuzzy set and the possibility distribution function is a relational assignment.

The values of t chosen to represent the possibility distribution function of failure time depend on the choice of the expert e.g. days, weeks, months, or years. Just as an expert can give an opinion on the possibility distribution of failure time of a generating unit, that expert can give an opinion on the possibility distribution of repair completion of a generating unit at various times. The time scales for failure time data and repair time data are usually quite different e.g. a generating unit might fail after working for months whereas the same unit can be repaired in weeks.

For the problem considered in the thesis, rf (t) is evaluated via subjective expert opinion which assigns fuzzy numbers to the failure possibility , decreasing from maximum 1 over t Є [0,5]. The ∏f (t) gives a fuzzy measure of the possibility of the unit failure at time ≤ t and such a measure differs from conventional probability.

The experts presume that as t increases, rf (t) decreases from maximum of 1. For a new unit the expert might give a possibility distribution by comparison with similar generating units under similar conditions.

$$\prod f(t) = \max (rf(u): u \in [0,t]) \tag{3.10}$$

The max operator has been used in aggregating the possibilities in above equation. However any other operator can be used for such aggregation. Kf(t) = 1 - ∏f (t). ∏r (t) is found in the same manner as ∏f (t), but by using the values of rr (t). Kr (t) = 1 - ∏r (t) .

$$\text{Fuzzy MTTF} = \sum Kf(t), \quad t \in N \tag{3.11}$$

The sum represents the area over Kf (t) divided by Kf (t) = 1 which is the height of Kf curve .

$$\text{Fuzzy MTTR} = \sum Kr(t), t \in N \tag{3.12}$$

$$\text{Fuzzy MTBF} = \text{Fuzzy MTTF} + \text{Fuzzy MTTR} \tag{3.13}$$

Fuzzy FOR is calculated as given in equation (3.14)

$$\text{Fuzzy FOR} = \text{Fuzzy MTTR} / \text{Fuzzy MTBF} \quad (3.14)$$

3.5.1 Case Study with Numerical Results

A captive power plant for an aluminum smelter, represented by Figure ------ is chosen for case studies having 6 thermal units, each of 120 MW and the optimum maintenance schedule obtained from levelized reserve method using hybrid GA/SA is given in Table 4.5. The effect of uncertainties related to forced outage rate (FOR) of generating units on the optimum schedule is investigated using the PROFUST model and the associated risk even with an optimum schedule is expressed as fuzzy loss of load probability (FLOLP), represented by Figures 8 and 9. The possibilistic fuzzy model assumes possibilistic approach for state behavior, instead of probabilistic state behavior, but retains the fuzzy state concept of PROFUST model.

The Fuzzy possibility distribution of MTTF is computed for periods following a particular maintenance to periods preceding the subsequent maintenance. The periods preceding nm=3 are the periods beginning of the planning horizon up to the maintenance nm=3. Since the MTTF decreases with progression of time, therefore the possibility distribution is considered to possess a maximum value of 1 at the beginning of the planning horizon and gradually decreases to 0.1 just before the maintenance nm=3. After the maintenance is carried out, the MTTF value increases, but with progression of time, it again decreases. But since nm=3 is a class S maintenance for maintenance of boiler and generator only, therefore the possibility immediately after such maintenance is 0.896 instead of 1. But the possibility distribution immediately after nm=4 is considered as 1 and the possibility distribution decreases with progression of time, till the next maintenance nm=5 is carried out. Similarly, the possibility distribution immediately after nm=5 is considered as 1 and decreases with progression of time. The possibility distribution increases with increase of time due to the fact that MTTR increases as time duration from previous maintenance increases.

The possibility distribution of MTTR increases as duration from previous maintenance increases, but again the value reduces towards zero as maintenance reduces the MTTR value.

Table 4. Comparison of FOR values obtained from PRUFUST and POSFUST models

Unit No	FOR using PROFUST model	FOR using POSFUST model
1	0.1111	0.1106
2	0.1008	0.1053
3	0.0906	0.0909
4	0.0704	0.0742
5	0.0603	0.0645
6	0.0402	0.0464

Possibility measures are directly connected with fuzzy sets via the associated possibility distribution functions in the POSFUST model. The possibility distribution represents the degree of membership for some linguistic variable, but the membership values are strictly monotonic as they are for an ordered possibility distribution. It is seen that in the POSFUST model, the FOR values are slightly greater than those values obtained from PROFUST model (except for unit 1) as the possibility distribution is nonprobabilistic. Thus, the PROFUST model is the approximation of the POSFUST model, which gives more accurate values of FOR when the observations pertaining to failure history of the generating units are a few. The comparisons of FLOLP for PROBIST, PROFUST and POSFUST models are represented graphically in Figures 10 and 11 respectively corresponding to cases when no unit is under maintenance and when one unit is under maintenance respectively.

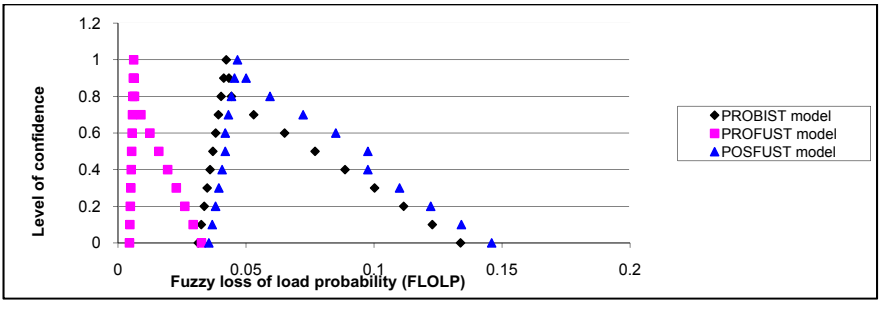


Fig. 10. Comparison of FLOLP values for PROBIST, PROFUST and POSFUST models when no unit under maintenance

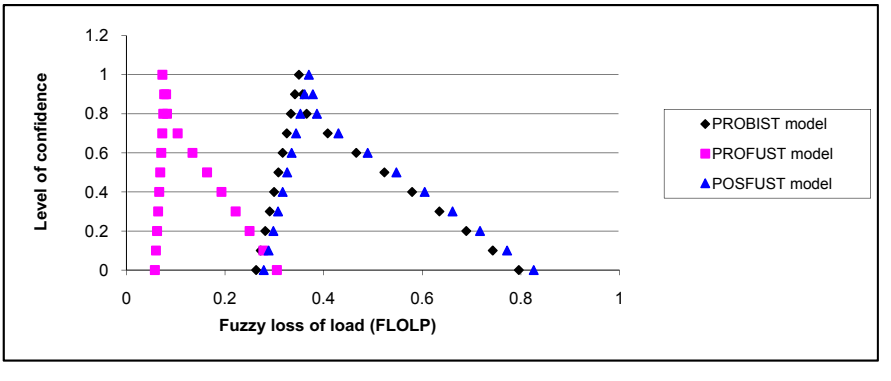


Fig. 11. Comparison of FLOLP values for PROBIST, PROFUST and POSFUST models when one unit under maintenance

When no unit is under maintenance and the effect of uncertainties related to FOR are considered in PROFUST and POSFUST models along with load forecasting uncertainties, the FLOLP has a range of 0.030761 to 0.131796 and 0.035519 to 0.145957 respectively for the maximum load uncertainty corresponding to level of confidence α having value of zero, as shown in Figure 10. But when the uncertainty related to FOR of generating units are ignored in case of PROBIST model, for the same value of α=0, the FLOLP has a range of 0.04501 to 0.032639 indicating lesser range of LOLP as uncertainties are reduced. Similarly, for α=1, the FLOP has a crisp values of 0.0415125 and 0.046701 respectively when the uncertainties related to FOR values are considered in PROFUST and POSFUST models. For the same case, the crisp value becomes 0.00616027 when the uncertainties related to FOR are ignored in the classical PROBIST model. Thus the FLOLP range denotes the risk limits associated with the uncertainties and it is seen that the more the uncertainty the greater is the range of FLOLP due to greater fuzziness. The similar trend is found when one unit is under maintenance as evidenced in Figure 11.

POSFUST model that uses possibility theory is a measure-theoretic counterpart of fuzzy set theory based upon the standard fuzzy operations. It provides us with appropriate tools for processing incomplete information expressed in terms of fuzzy propositions and consequently plays a major role in fuzzy logic. Both probability and possibility measure are uniquely represented by distribution function, but their normalization requirements are very different. Values of each probability distribution are required to add to 1, while for possibility distributions the largest values are required to be 1. Moreover, the later requirement may even be abandoned when possibility theory is formulated in terms of fuzzy sets.

4 Conclusion

The fuzzy reliability evaluations for power system are of immense importance because of the uncertainty related to load demand and available history of failure of components can be taken care using fuzzy models in some suitable way. Some representative cases in this chapter prove the efficacy of the fuzzy reliability analysis using such models. The following are some of the salient features related to these fuzzy models.

4.1 Type I Fuzzy Reliability Analysis for Incorporating Load Forecasting Uncertainties

Type I fuzzy calculations at the hierarchical level 1(HL-1) are related to the uncertainty in defining the power consumption and therefore measuring the impacts of failure of supply. Although the relation between a changing load and a varying calculated index is far from linear, but still its characteristics are well captured by the fuzzy load description in the reported research.

The load forecasting uncertainties incorporated through the fuzzy model corroborate the fact that a small uncertainty, represented by a linear variation of the predicted load, has an ample effect on the reliability index uncertainties. Also it is evident that although the relation between predicted load and a varying FLOLP is not linear but the model gives some boundary of variation. Since captive power plants cater to power sensitive utilities, the boundary of variation of reliability indices is very helpful for user industries to assess risks when the load requirements are not properly defined.

4.2 Probabilistic Fuzzy State Model

In many cases, the classical probabilistic binary state (PROBIST) model representing binary state reliability indices such as forced outage rate (FOR) by crisp number seems to be inadequate from pragmatic perspective. Most of the reliability data are obtained from databases. But the data associated with equipment may not be exactly that exists in the database; either because it was not installed under the same conditions or just some new types of equipment are in vogue. Consequently some uncertainties are associated with component indices due to lack of upgradation of data. A probabilistic fuzzy state (PROFUST) model takes into account the uncertainties related to forced outage rate by considering the uncertainties related to mean time to failure (MTTF) and mean time to repair (MTTR) for the generating units.

The PROFUST model seems to be more promising compared to the classical PROBIST model for reliability analysis based on system adequacy for captive power plant maintenance scheduling as the uncertainties related to FOR are well captured in fuzzy model incorporating expert opinions. The additional impact of the uncertainties in the load represented by fuzzy load model has also been incorporated in hybrid fuzzy model to calculate FLOLP. The reliability index FLOLP reflects uncertainties in terms of quantified risk and this is of immense use for planning related to captive power plants because the degree of acceptable risk for a captive power plant is implicitly related to the requirements of the process industry that is serving as the load for the captive power plant.

The advantage of using the probabilistic fuzzy model lies in the fact that it is not exactly a replacement of the old faithful classical probabilistic model, but its enhancement for better engineering approximation of FOR of generating units by reflecting the impact of aging of units as well as the type of maintenance for a realistic schedule.

4.3 Possibilistic Fuzzy State Model

The standard approaches of reliability engineering rely on the probability model which is often less appropriate for this task. Probability based analyses usually require more information about the system than is known. Fuzzy logic offers an alternative to the probability paradigm through the use of possibility paradigm. The possibilistic approach for the system behavior allows for quantitative

reliability evaluations that preserve the uncertainty in original data. The possibility model deals with uncertainty in a way that avoids making unwarranted conclusions and makes the consequences of the required assumptions more clear.

Possibility measures are directly connected with fuzzy sets via the associated possibility distribution functions in the POSFUST model. The possibility distribution represents the degree of membership for some linguistic variable but the membership values are strictly monotonic as they are for an ordered possibility distribution. Usually PROFUST model adopts more conservative approach about its failure, but POSFUST model gives a realistic value which seems to be less compared to those obtained from PROFUST model. Thus the PROFUST model is the approximation of the POSFUST model. However, it gives more accurate values of FOR when the observations pertaining to failure history of the generating units are a few.

References

1. Billinton, R., Allan, R.N.: Reliability Evaluation of Power System. Plenum Press, New York (1984)
2. Billinton, R.: Power System Reliability Evaluation. Gordon and Breach Publishers, New York (1974)
3. Endrenyi, J.: Reliability Modeling in Electric Power Systems, pp. 137–145. John Wiley and Sons, Chichester (1980)
4. EI-Hawary, M.E. (ed.): Electric Power Applications of Fuzzy System. IEEE Press, New York (1998)
5. Mohanta, D.K., Sadhu, P.K., Chakrabarti, R.: Fuzzy Reliability Evaluation of Captive Power Plant Maintenance Scheduling Incorporating Uncertain Forced Outage Rate and Load Representation. Electric Power System Research 72(1), 73–84 (2004)
6. Mohanta, D.K., Sadhu, P.K., Chakrabarti, R.: Fuzzy Markov Model for Determination of Generating Unit Fuzzy State Probabilities Including the Effect of Maintenance Scheduling. IEEE Transactions on Power Systems 20(4), 2117–2126 (2005)
7. Mohanta, D.K., Khaitan, S.K., Sadhu, P.K., Chakrabarti, R.: Emerging Trends in Fuzzy Based Power System Reliability Analysis. In: The Proceedings of International Conference on Emerging Technologies, December 19-21. KIIT, Bhubaneswar (2003)
8. Chakrabarti, R., Mohanta, D.K., Sadhu, P.K.: Possibilistic Approach for Evaluation of Forced Outage Rates of Generating Units Including the Effect of Maintenance Scheduling. Electrical Journal of the Institution of Engineers (India) 87, 48–52 (2007)
9. Miranda, V.: Fuzzy Reliability Analysis of Power Systems. In: Proceedings of the Twelfth Power System Computation Conference, PSCC 1996, Dresden, Germany, August 1996, vol. 1, pp. 558–566 (1996)
10. Saraiva, J.T., Miranda, V., Pinto, L.M.V.G.: Impact on Some Planning Decision from a Fuzzy Modeling of Power Systems. IEEE Transactions on Power Systems 9(2) (1994)
11. Momoh, J.A., Ma, X.W., Tomsovic, K.: Overview and Literature Survey of Fuzzy Set Theory in Power Systems. IEEE Transactions on Power Systems 10(3), 1676–1690 (1995)

12. Fayyaz Noor, S., Mc.Donald, J.R.: Forced Outage Rates of Generating Units Based on Expert Evaluation. IEEE Transactions on Reliability 45(1) (March 1996)
13. Zadeh, L.: Fuzzy Sets as a Basis for a Theory of Possibility. Fuzzy Sets and Systems 1, 3–28 (1978)
14. Cai, K.Y.: Fuzzy Variables as a Basic for a Theory of Fuzzy Reliability in Possibility Context. Fuzzy Sets and Systems 42, 145–172 (1991)
15. Ross, Timothy, J.: Fuzzy Logic With Engineering Applications. McGraw Hill, Inc., New York (1995)
16. Klir, G., Folger, T.: Fuzzy Sets, Uncertainty and Information. Prentice Hall, Englewood Cliffs (1998)

Load Forecasting and Neural Networks: A Prediction Interval-Based Perspective

Abbas Khosravi, Saeid Nahavandi, and Doug Creighton

Abstract. Successfully determining competitive optimal schedules for electricity generation intimately hinges on the forecasts of loads. The nonstationarity and high volatility of loads make their accurate prediction somewhat problematic. Presence of uncertainty in data also significantly degrades accuracy of point predictions produced by deterministic load forecasting models. Therefore, operation planning utilizing these predictions will be unreliable. This paper aims at developing prediction intervals rather than producing exact point prediction. Prediction intervals are theatrically more reliable and practical than predicted values. The delta and Bayesian techniques for constructing prediction intervals for forecasted loads are implemented here. To objectively and comprehensively assess quality of constructed prediction intervals, a new index based on length and coverage probability of prediction intervals is developed. In experiments with real data, and through calculation of global statistics, it is shown that neural network point prediction performance is unreliable. In contrast, prediction intervals developed using the delta and Bayesian techniques are satisfactorily narrow, with a high coverage probability.

1 Introduction

Generating optimal schedules for power systems can potentially save millions of dollars per year for utility companies [1]. To remain competitive in the privatized and deregulated markets of power generation, it is vital for companies to reduce their operation cost, through formulating optimal schedules for generating power. Effectiveness of electricity power generation plans significantly depends on the accuracy of the future load predictions. Both over-forecasting and under-forecasting may result in financially disastrous circumstances costing millions of dollars for big utility companies.

Owing to the significant impacts of future demand prediction on reliability of energy management plans, a variety of methods have been proposed in scientific and practical literature for precise load forecasting. Broadly speaking, these techniques

Abbas Khosravi, Saeid Nahavandi, and Doug Creighton
Centre for Intelligent Systems Research (CISR), Deakin University, Vic, 3217, Australia
E-mail: {akhos,saeid.nahavandi,douglas.creighton}@deakin.edu.au

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 131–150.
springerlink.com

Table 1. Nomenclature and Abbreviations

x_i	i-th nonrandom m-input vector
D_{train}	Training datasets
D_{test}	Test dataset
Θ	Original set of neural network parameters
Θ^*	The true set of NN parameters
$\widehat{\Theta}$	The set of NN parameters obtained using least square technique
$\widehat{y_i}$	Neural network scalar output
Y	Vector of outputs
$\Phi(x_i, \Theta^*)$	Nonlinear function of true regression
ε_i	i-th error term associated with modeling function
σ^2	Variance of ε_i
$\bigtriangledown^T_{\Theta^*}\hat{y}$	Gradient of neural network output with respect to its set of parameters (Θ^*)
J	Jacobian matrix of neural network
I	Identity matrix
s	Standard deviation estimate
λ	Regularizer factor
$t_d^{1-\frac{\alpha}{2}}$	$1-\frac{\alpha}{2}$ quantile of a cumulative t-distribution function with d degrees of freedom
M	Particular neural network model
ρ and β	Hyperparameters in Bayesian technique
Θ^{MP}	Neural network parameters computed using Bayesian learning algorithm
H^{MP}	Hessian matrix computed for Θ^{MP}
$z^{1-\frac{\alpha}{2}}$	$1-\frac{\alpha}{2}$ quantile of a normal distribution function
$L(x_i)$ & $U(x_i)$	Lower and upper bounds of PI for sample x_i
ξ	Range of the underlying target
$\sigma(\cdot)$	Sigmoidal function
η & μ	Parameters of the sigmoidal function
n_i^{Upper}	Upper bound for the number of neurons in the i-th layer of neural network model
MSE	Mean Squared Error
MAPE	Mean Absolute Percentage Error
PICP	Prediction Interval Coverage Probability
NMPIL	Normalized Mean Prediction Interval Length
CLC	Coverage Length-based Criterion

can be divided into two categories: parametric models, and artificial intelligence-based techniques (non-parametric techniques) [1] [2]. Regression models, Kalman filter, and time-series prediction methods are among the parametric techniques that have been widely applied for load forecasting [3] [4] [5]. Despite their easy implementation, they have limited power in finding and modeling nonlinear dependencies amongst independent and dependent variables. Furthermore, they require a priori model needs to be assumed or a priori assumptions on the properties of data to be made.

It has been shown that artificial intelligence-based methods, such as Neural Networks (NNs) outperform their traditional rivals, not only in the area of load forecasting [1] , but also in many other engineering and science domains. Paliwal et al. [6] comprehensively reviewed comparative studies on applications of NNs in accounting and finance, health and medicine, engineering, manufacturing, and marketing.

After reviewing around 100 comparative studies, they concluded that in the majority of cases, no matter what the type of data or from which domain it comes, NN models outperform their traditional rivals.

The established reputation of NN in modeling and analysis of complex systems is due to their nonparametric, adaptive, and noise tolerant properties. As biologically inspired analytical techniques, they have capability of learning and modeling very complex nonlinear relationships. Theoretically, multilayered feed forward NNs are universal approximators, and as such, have excellent capability of approximating any nonlinear mapping to any degree of accuracy [7]. Their generalization to the not-yet-observed data is better than other techniques. Furthermore, they require no a priori assumptions on the properties of data to be made. NNs have been widely employed, among others, for modeling, prediction, classification, optimization, and control purposes [8] [9] [10] [11].

Like other areas of science and engineering, reports on successful applications of NNs for short/medium/long term load forecasting are numerous [1] [2] [12] [13] [14] [15]. In some cases, fuzzy systems have been integrated into NNs for improving prediction results [16] [17]. In the majority of studies and research conducted so far in the area of load forecasting, focus has been on developing NN models for point prediction, without any indication of its likely accuracy. One should make note of the fact that loads often show completely nonlinear and in some cases chaotic behaviors. Their fluctuations through the time are erratic and influenced by many factors. These factors may totally or partially be unknown for modelers. Even known, often information about them is uncertain (e.g., temperature and weather conditions in the upcoming hours or days). Accompanied with uncertainty are probabilistic events whose occurrence may result in multiple realities for the future of a system. With these features, power systems resemble volatile systems whose outcome (load demand in future) depends on many uncertain and unknown factors. A point that often has been ignored in the conducted studies in literature is that NNs are deterministic models and their applications for predicting the future of a stochastic system is always questionable [18]. Uncertainties and probabilistic events highly contribute to the degradation of performance of NN models for load forecasting. Negative consequences raised from the stochastic nature of power systems cannot be compensated solely through increasing NN size (neither hidden layers nor neurons) or repeating its training procedure. As NN load forecasters offer a point estimate without any measure of its accuracy, making decisions based on such predictions may spell disaster for different components of a power system. Load forecasting literature is void of information about incorporating uncertainties of power systems into point prediction of system future. Part of the literature has materialized into a series of studies whose objective has been to adopt NNs for prediction of stochastic systems future [18] [19]. Construction of Prediction Intervals (PIs) has been proposed in literature to remedy problems of NNs. Mathematically, a prediction interval with confidence level of $(1-\alpha)\%$ is a random interval developed based on past observations $x = (x_1, x_2, \cdots, x_n)$ and built for unseen observations $X_{n+h}, h \geq 1$,

$$PI = [L(x), U(x)] \tag{1}$$

such that $P(L(x) < x_{n+1} < U(x)) = 1 - \alpha$. Without construction of prediction intervals, the validity of decisions made based on point prediction is always questionable. When developing theories for constructing PIs, different sources of uncertainty are considered. They are practically more useful than confidence intervals and more reliable than predicted points. PIs also carry much more meaningful information than point predictions: they have an indication of their accuracy (level of confidence), they are constructed after considering and modeling different sources of uncertainty, and finally, they provide a range of probable futures.

In literature, there are a couple of techniques for creating prediction/confidence intervals for NNs: (a) the delta method, which is based on a Taylor expansion of the regression function [20] [21], (b) the bootstrap technique which is essentially a resampling method [22], and (c) the Bayesian approach, based on Bayesian statistics [23]. Selection of any of these techniques for constructing PIs depends on problem domain, computation burden, number of available samples, and analysis purpose. Comparative studies on useability and practicality of these techniques have been carried out in literature [24] [25] [26].

Although theories behind PI construction using NNs date back to 90s, their application has only recently proliferated. Many have used them instead of point predictions in different engineering and science fields, among others, including temperature prediction [27], boring process prediction [28], paper curl prediction [25], watershed simulation [29], modeling of solder paste deposition process [30], baggage handling system [31] [32], and time series forecasting [33].

The main focus of this research is on the delta and Bayesian techniques for PI construction [23] [27] [28] [30] [34]. Although these techniques have been used in other areas of engineering, there is no report on their applications in the area of power engineering, and in particular, load forecasting. A review of the literature reveals only a limited body of knowledge on PI construction in the area of load forecasting. As loads many times show volatile behavior, delta and Bayesian techniques can be employed to enhance the reliability of prediction task, and therefore, the practicality of operation schedules.

In literature many error-related measures, such as Mean Absolute Percentage Error (MAPE) and Mean Squared Error (MSE), have been proposed for assessment of point predictions. Needless to say, such measures are not applicable to PIs. To the best of our knowledge, the main focus of literature is on construction of PIs, without any objective assessment of their quality in terms of length and coverage probability [28] [25] [30]. Often the coverage probability of PIs is the main concern and PIs are judged on the basis of closeness of this index to the nominal confidence level of PIs ($(1-\alpha)\%$) [25] [35]. One often ignored measure related to the quality of PIs is length. In this paper, a new assessment measure is developed for PIs for objective assessment of PIs constructed using different techniques. The proposed index is composed of both coverage probability and length of PIs. Application of this new index makes possible to objectively select NN forecaster that yield narrower PIs with higher coverage probability.

For the first time in the area of power engineering, delta and Bayesian techniques are used for constricting PIs for outputs of NNs in a load forecasting problem.

Constructed PIs are then assessed and judged based on the proposed evaluation index. In the experiments with real data, it is demonstrated that point prediction error is always big and, therefore, results are unreliable. In contrast, constructed PIs using delta and Bayesian techniques are satisfactorily narrow with a high coverage probability close to the nominal confidence level.

The rest of this paper is organized as follows: Section 2 briefly introduces the delta and Bayesian methods for constructing PIs. In Section 3, a new measure for comprehensive evaluation of PIs is developed. In Section 4, experimental results for constricting PIs for a load forecasting problem are represented and discusses. Section 5 concludes the paper with some remarks and directions for future research in this domain.

2 Theory and Background

This section provides a brief overview of the delta and Bayesian techniques for constructing PIs. It is assumed that the intended audience of this paper are familiar with NNs, their training techniques, and issues related to them such as over-fitting or under-fitting. Detailed description and mathematical discussion about different types of NNs can be found in [23].

Both Bayesian and delta techniques are adopted for constructing PIs for outputs of NNs. Before going further, it is necessary first to distinguish the difference between confidence intervals and prediction intervals. Confidence intervals provide a measure of the uncertainty between the prediction and the expected mean value of the dependent variables. On the other hand, PIs deal with the accuracy of prediction of an individual output. Based on this discussion, it is obvious that prediction intervals are wider than confidence intervals and cover them [36].

2.1 Delta Technique

The delta technique is based on representing and interpreting NNs as nonlinear regression models. This allows applying standard asymptotic theory to them for constructing prediction intervals. According to this, one may represent them as below,

$$y_i = \Phi(x_i, \Theta^*) + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{2}$$

where n is the number of samples. The term ε_i is the error associated with the modeling function and its misspecification. $\hat{\Theta}$, an estimate of Θ^*, is obtained through minimization of sum of squared error (SSE) cost function,

$$SSE = \sum_{i=1}^{n} \left(y_i - \Phi(x_i, \widehat{\Theta}) \right)^2 \tag{3}$$

where prediction for the i-th sample is obtained using $\hat{\Theta}$,

$$\hat{y}_i = \Phi(x_i, \hat{\Theta}) \tag{4}$$

A first-order Taylor expansion of (2) around the true values of model parameters (Θ^*) can be expressed as,

$$\hat{y}_i = \Phi(x_i, \Theta^*) + \bigtriangledown^T_{\Theta^*} \hat{y}_i (\hat{\Theta} - \Theta^*), \quad i = 1, 2, \ldots, n \tag{5}$$

Dimensionality of $\bigtriangledown^T_{\Theta^*} \hat{y}_i$ in (5) is $1 \times p$. With the assumption that ε_i in (2) are independently and normally distributed $N(0, \sigma^2)$, the $(1-\alpha)\%$ PI for y_i is [20],

$$\hat{y}_i \pm t_d^{1-\frac{\alpha}{2}} s \sqrt{1 + \bigtriangledown^T_{\Theta^*} \hat{y}_i (J^T J)^{-1} \bigtriangledown_{\Theta^*} \hat{y}_i} \tag{6}$$

where $\bigtriangledown^T_{\Theta^*} \hat{y}_i$ and J are defined as follows,

$$\bigtriangledown^T_{\Theta^*} \hat{y}_i = \left[\frac{\partial \Phi(x_i, \Theta^*)}{\partial \Theta^*_j} \right]_j, \quad j = 1, 2, \ldots, p \tag{7}$$

$$J = \left[\frac{\partial \Phi(x_i, \Theta^*)}{\partial \Theta^*_j} \right]_{i,j}, \quad i = 1, 2, \ldots, n, j = 1, 2, \ldots, p \tag{8}$$

The delta method is based on minimization of the quadratic cost function in (3). This cost function is only related to the prediction errors and does not put any penalties on the network size or does not constraint the parameter magnitudes. This may result in singularity of matrix $J^T J$, that in turn makes computed PIs less reliable. To remedy this problem, one alternative is to include some regularizing terms in (3). Based on the weight decay concept [23], De Veaux et al. [21] have added the sum of squares of adaptive parameters to the cost function in (3),

$$\left(Y - \hat{Y}\right)^T \left(Y - \hat{Y}\right) + \lambda \hat{\Theta}^T \hat{\Theta} \tag{9}$$

where λ is the regularizing factor [23]. The generalization power of NN models is highly improved through using (9) as the training cost function. Rebuilding PIs based on (9) will yield the following PIs,

$$\hat{y}_i \pm t_d^{1-\frac{\alpha}{2}} s \sqrt{1 + \bigtriangledown^T_{\Theta^*} \hat{y}_i (J^T J + \lambda I)^{-1} (J^T J) (J^T J + \lambda I)^{-1} \bigtriangledown_{\Theta^*} \hat{y}_i} \tag{10}$$

The procedure for calculation of s in (10) has small differences with the one used in (6). These differences arise from the corresponding cost function considered in each method. An acceptable estimate of s in (10) has been given below,

$$s = \frac{\sqrt{SSE}}{n - trace(2\Gamma - \Gamma^2)} \tag{11}$$

where,

$$\Gamma = J(J^T J + \lambda I)^{-1} J^T \tag{12}$$

Generally, training NNs based on (9) and then using (10) for constructing PIs yields more acceptable results. In case of using (3) and (6) for analysis, some trained networks must be discarded due to quick convergence of optimization to some values, which are not close to the minimum. More discussion on these issues can be found in [21].

2.2 Bayesian Technique

Bayesian technique for constructing PIs is based on NN Bayesian training algorithm [23]. The method aims at developing a NN with the best performance and the least parameters so that the problem of over-fitting can be effectively controlled. The Bayesian approach has the considerable benefits that validation and test sets are not needed for NN training. From a Bayesian inference perspective, the NN parameters are considered as random variables with unknown distributions. The posterior probability of a NN model with the set of parameter Θ and the training dataset D can be determined using Bayes' rule:

$$P(\Theta|D,\rho,\beta,M) = \frac{P(D|\Theta,\beta,M)P(\Theta|\rho,M)}{P(D|\rho,\beta,M)} \tag{13}$$

where $P(D|\Theta,\beta,M)$ and $P(\Theta|\rho,M)$ are the likelihood function of data occurrence and the prior density of parameters respectively. $P(D|\rho,\beta,M)$ is also a normalization factor enforcing that total probability is one. It is possible to consider different distributions for $P(D|\Theta,\beta,M)$ and $P(\Theta|\rho,M)$. In order to simplify further analysis, it is often assumed that they are Gaussian,

$$P(D|\Theta,\beta,M) = \frac{1}{Z_D(\beta)} e^{-\beta E_D} \tag{14}$$

and

$$P(\Theta|\rho,M) = \frac{1}{Z_\Theta(\rho)} e^{-\rho E_\Theta} \tag{15}$$

where $Z_D(\beta) = \left(\frac{\pi}{\beta}\right)^{\frac{n}{2}}$ and $Z_\Theta(\rho) = \left(\frac{\pi}{\rho}\right)^{\frac{p}{2}}$. n and p are number of training samples and NN parameters respectively. Substituting (14) and (15) into (13) results in,

$$P(\Theta|D,\rho,\beta,M) = \frac{1}{Z_F(\beta,\rho)} e^{-(\rho E_\Theta + \beta E_D)} \tag{16}$$

When training NNs using Bayesian method, the purpose is to maximize $P(\Theta|D,\rho,\beta,M)$. Such maximization can be achieved through minimization of the following cost function,

$$E(\Theta) = \rho E_\Theta + \beta E_D \tag{17}$$

where E_D is SSE defined in (3) and E_Θ is the sum of squares of the network weights ($\Theta^T\Theta$). It has been shown [23] [37] that the following values for β and ρ maximize the posterior probability function in (16),

$$\beta^{MP} = \frac{\gamma}{E_D(\Theta^{MP})} \tag{18}$$

$$\rho^{MP} = \frac{n-\gamma}{E_\Theta(\Theta^{MP})} \tag{19}$$

γ is the effective number of parameters,

$$\gamma = p - 2\rho^{MP} tr\left(H^{MP}\right)^{-1} \tag{20}$$

where H^{MP} is the hessian matrix of $E(\Theta)$ computed for Θ^* ($H^{MP} = \rho\nabla^2 E_\Theta + \beta\nabla^2 E_D$). γ is a measure of the number of NN parameters effectively used for error function minimization. Usually, the LevenbergMarquardt optimization algorithm is applied to approximate the Hessian matrix (H^{MP}) [38].

Application of this technique for training NNs results in NNs that the variance of their prediction is,

$$\sigma_i^2 = \sigma_D^2 + \sigma_{\Theta^{MP}}^2 = \frac{1}{\beta} + \nabla_{\Theta^{MP}}^T \hat{y}_i \ (H^{MP})^{-1} \ \nabla_{\Theta^{MP}} \hat{y}_i \tag{21}$$

Therefore, a $(1-\alpha)\%$ PI for the i-th future sample can be constructed as follows,

$$\hat{y}_i \pm z^{1-\frac{\alpha}{2}} \left(\frac{1}{\beta} + \nabla_{\Theta^{MP}}^T \hat{y}_i \ (H^{MP})^{-1} \ \nabla_{\Theta^{MP}} \hat{y}_i\right)^{\frac{1}{2}} \tag{22}$$

Further information about this technique and its detailed mathematical discussion can be found in [23] [37].

3 A New Assessment Measure for Prediction Intervals

As discussed before, literature does not offer a suitable measure for comprehensive assessment of PIs. In this section a new general examination measure is proposed which covers both important aspects of PIs: length and coverage probability. As the proposed measure is general and developed based on features of PIs (not the utilized method for constructing PIs), it can be applied in other relevant studies as well.

Theoretically, one can characterize PIs based on their length and coverage probability. One approach for quantitative assessment of PI lengths is to normalize each interval length with regard to range of targets. Following this, a measure called Normalized Mean Prediction Interval Length (NMPIL) can be obtained as follows:

$$NMPIL = \frac{1}{n\xi} \sum_{i=1}^{n} (U(x_i) - L(x_i)) \tag{23}$$

Normalization of PI length by the range of targets makes the objective comparison of PIs possible, regardless of techniques used for their construction or magnitudes of the underlying targets. The upper bound of NMPIL is one, obtained for the case that minimum and maximum of targets are considered as upper and lower bounds

of PIs for all targets. Usually, the smaller the NMPIL, the more useful the PIs. The lower bound of NMPIL is model dependent and is dominated by MSE of NN modes. Assuming that in the ideal case, the gradient term in (10) vanishes for unobserved samples, one can obtain the lower bound of NMPIL for the delta technique as follows,

$$NMPIL_{minD} = \frac{1}{n\xi} \sum_{i=1}^{n} 2t_d^{1-\frac{\alpha}{2}} s \tag{24}$$

If the Bayesian method is applied, this lower bounds will be,

$$NMPIL_{minB} = \frac{1}{n\xi} \sum_{i=1}^{n} 2z^{1-\frac{\alpha}{2}} \beta^{\frac{-1}{2}} \tag{25}$$

Practically, achieving $NMPIL_{min}$ for PIs is far remote. This stems from the fact that gradient terms in (10) and (22) are not ignorable. Indeed, they are often big for unobserved (test) samples, as test samples are not used in the training stage of NNs.

While NMPIL relates to the length of PIs, another measure is required for monitoring coverage of PIs. If PIs are deliberately squeezed in favor of achieving smaller NMPIL, many targets may drop out of PIs. Therefore, another measure is required for quantification of this phenomenon. The PI Coverage Probability (PICP) indicates the probability that the underlying targets will lie within the constructed PIs. It can be calculated through counting the covered targets by PIs:

$$PICP = \frac{1}{n} \sum_{i=1}^{n} c_i \tag{26}$$

where $c_i = 1$ if $y_i \in [L(x_i), U(x_i)]$, otherwise $c_i = 0$. Theoretically, PICP should be as close as possible to its nominal value, $(1-\alpha)\%$, the confidence level that PIs have been constructed based on. Unfortunately, in reality this often does not happen. Imperfectness of PICP is attributable to the presence of noise in samples and severe effects of uncertainty. Other issues such as under-fitting and over-fitting (which are direct results of using (very) small or big NNs) also contribute to the unsatisfactory smallness of PICP.

It is always desirable to construct PIs whose PICP is the highest possible value. Such high PICP can be simply achieved through considering target ranges as PIs for all samples. Needless to say, wide PIs like these ones are practically useless. This argument makes clear that judgment about PIs based on PICP without considering length of PIs (here, NMPIL) is always subjective and biased. It is essential to evaluate PIs simultaneously based on their both key measures: length (NMPIL) and coverage probability (PICP).

Generally, PI lengths and PICP have a direct relationship. The wider the PIs, the higher the corresponding PICP. This means that as soon as PIs are squeezed, some targets will lie out of PIs, which results in a lower PICP. According to this discussion, the following Coverage-Length-based Criterion (CLC) is proposed for comprehensive evaluation of PIs in term of their coverage probability and lengths,

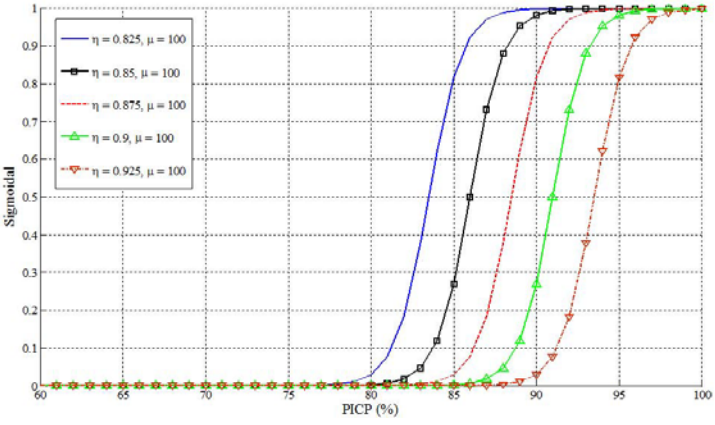


Fig. 1. Evolution of the sigmoidal function for different values of its parameters (η and μ)

$$CLC = \frac{NMPIL}{\sigma(PICP, \eta, \mu)} \tag{27}$$

where $\sigma(\cdot)$ is the sigmoidal function defined as follows,

$$\sigma(PICP, \eta, \mu) = \frac{1}{1 + e^{\eta(PICP-\mu)}} \tag{28}$$

The level of confidence that PIs have been constructed based on can be appropriately used as a guide for selecting hyperparameters of CLC. One reasonable principle is that we highly penalize PIs that their PICP is less than $(1-\alpha)\%$. This is based on the theory that the coverage probability of PIs in an infinite number of replicates will approach towards $(1-\alpha)\%$.

Fig. 1 demonstrates $\sigma(PICP, \eta, \mu)$ for different values of η and μ. It can be seen that the sigmoidal function sharply drops immediately after some values of PICP. These values are determined based on the confidence level of PIs, $(1-\alpha)\%$. According to curves in Fig. 1, if PICP is less than some nominal thresholds, CLC will highly increase, no matter what the length of PIs is. In this way, PIs with not satisfactorily high coverage probability are highly penalized. Generally, smallness of CLC is an indication of goodness of constructed PIs (simultaneously achieving small NMPIL and high PICP). Smallness or bigness of CLC is totally case-dependant. However, if PICP is sufficiently high, CLC and NMPIL will be almost the same.

4 Experiment and Numerical Analysis

4.1 Data and Experiment Procedure

The delta and Bayesian techniques are here applied for constructing PIs for load datasets used in [16]. The purpose of load forecasting is to predict the two-day-ahead

loads based on the past observations. The dataset includes records of consumed loads, weather condition, and calendar information for five years. For this case study, NN model includes 15 inputs and one output (load demand in the next day). Inputs are loads in the last hours (3), temperatures in three major cities in the last hours (9), day and month (2), and the corresponding load last year in the same day and hour (1). It is important to mention here that the main purpose of this paper is not to develop optimal NN models for load forecasting. Generally, accuracy of load forecasts depends on many factors including, among others, set of inputs (independent variables), NN structure, training algorithm (over-fitted or under-fitted), and targets' behavior. Often some kinds of analysis are conducted for selection of appropriate inputs or determining the optimal structure of NN models [39] [40]. The main objective here is to show that point prediction error always exists. Therefore, point predictions are unreliable. In this study, the aforementioned inputs are used for developing NN models based on the procedure schematically shown in Fig. 2. As NNs are data-driven techniques, their prediction performance highly depends on data, their initial parameters, and their structure. To avoid any subjective judgment about NN prediction capabilities, samples are split into training (70%) and test (30%) samples. Then NN is trained using training sets and examined using test samples. This procedure is repeated B times. Then average the results for different measures, including global statistics and CLC, are computed. Experiment continues until exploration of all potential structures. Through this way, effects of random initialization of NN parameters or data distribution are minimized.

In the experiments, different structures are examined for developing NN models. The two layer feedforward NN models are developed in the experiments. The quantity of neurons in their first and second hidden layers are changed within a grid $([1, n_i^{Upper}], i = 1,2)$. n_i^{Upper} can be determined based on the number of available samples and the proposed method in [31], guaranteeing $d \geq 30$. Over a grid including of $\prod_{i=1}^{2} n_i^{Upper}$ nodes, networks have been trained B times (totally $B \prod_{i=1}^{2} n_i^{Upper}$ NNs are investigated in experiment).

Table 2 lists some parameters and quantities used in both case studies. In order to give all independent variables equal chance to contribute to the built models, all variables are pre-processed to have zero mean and unit variance.

Table 2. Parameter values used in experiments

α	0.1
η	200
μ	0.875
B	5
n_1^{Upper}	10
n_2^{Upper}	10
D_{train}	70% of all samples
D_{test}	30% of all samples

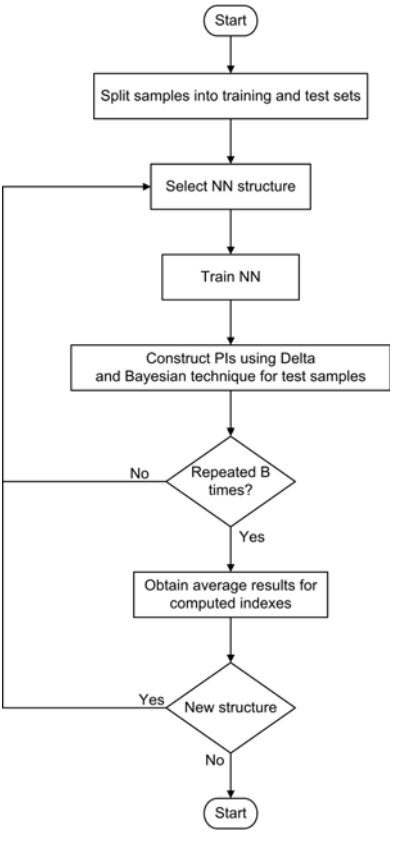


Fig. 2. Experiment procedure

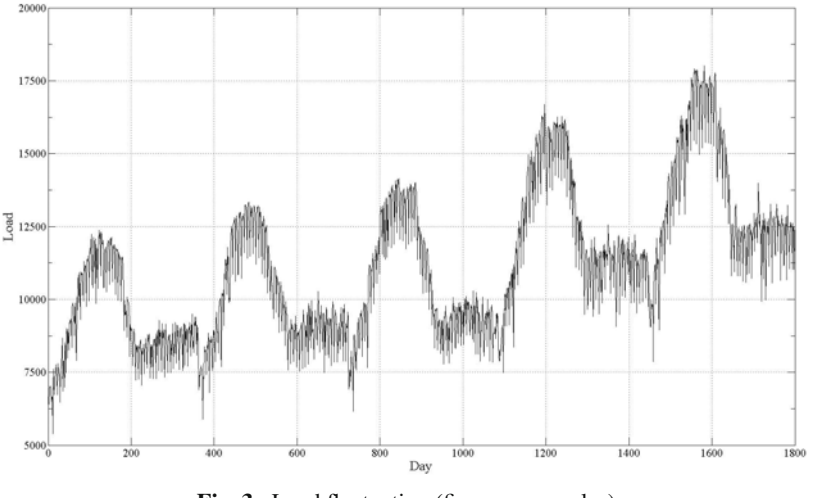


Fig. 3. Load fluctuation (five year samples)

4.2 Results and Discussion

The samples for experiments in this study have been plotted in Fig. 3. Visual inspection of samples illustrates that load consumption peaks happen in summer. Also, there is an upward trend in samples, which indicates growth of load consumption through the five years.

Fig. 4 shows the average coefficient of determination (R^2) for test samples computed over a gird of NNs. The maximum of R^2 is 94.8%, which is low and unsatisfactory. It is achieved for one of the biggest investigated NNs with 8 neurons in the first layer and 10 neurons in the second layer. Also the mean and standard division of R^2 are 91.8% and 0.84%. Taking into account these values and the maximum of R^2, one may conclude that no matter which structure is used or what the initial parameters are, NN point prediction performance is poor. A stable pattern in the obtained results is that as the NN size increases, the R^2 tends to decrease very slowly. Based on Fig. 4, this reduction plagues for NNs with more than 5 neurons in their second layer.

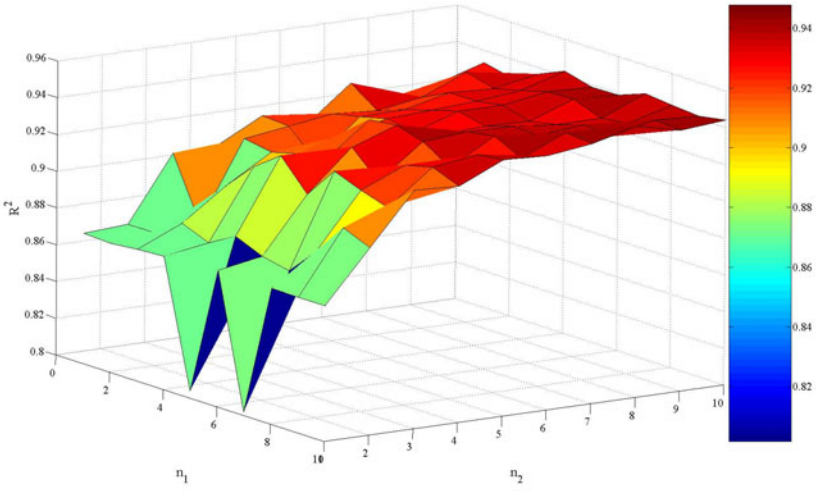


Fig. 4. Average R^2 for test samples over a grid of 100 NNs

The scatter plot of actual loads against the predicted ones has been given in Fig. 5. Prediction has been done using a NN with 8 neurons in its first layer and 5 neurones in its second layer. R^2 index for this particular case is 0.8513 that is again low. Examination of MAPE for this case shows that it is also quite big (227%). Even if the outliers are deleted from results, it does not become less than 75%.

Fig. 6 and Fig. 7 represent CLCs of PIs constructed using delta and Bayesian techniques, respectively. With the sake of better graphical representation, CLCs of a few outliers have been changed to 100. Both methods show promising results in term of length and coverage probability of PIs. Since CLC is always less than 100 (with the exception of a few outliers when using delta technique), it is reasonable to conclude that

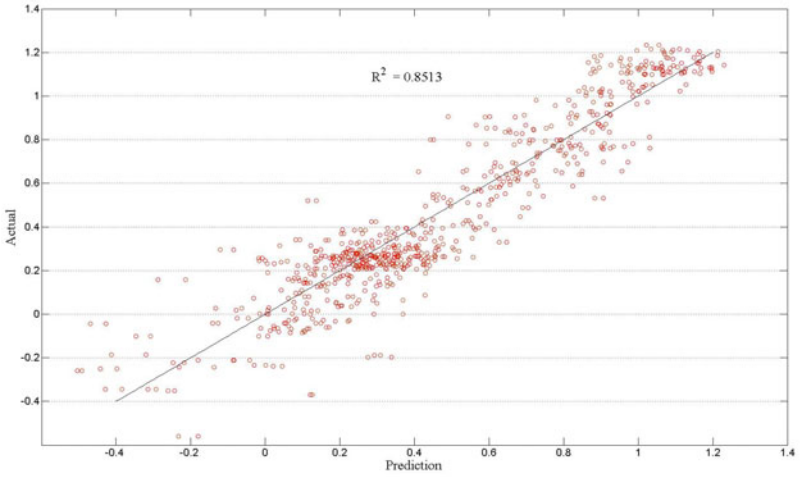


Fig. 5. Scatter plots between actual and forecasted loads using a NN with (8-5) structure

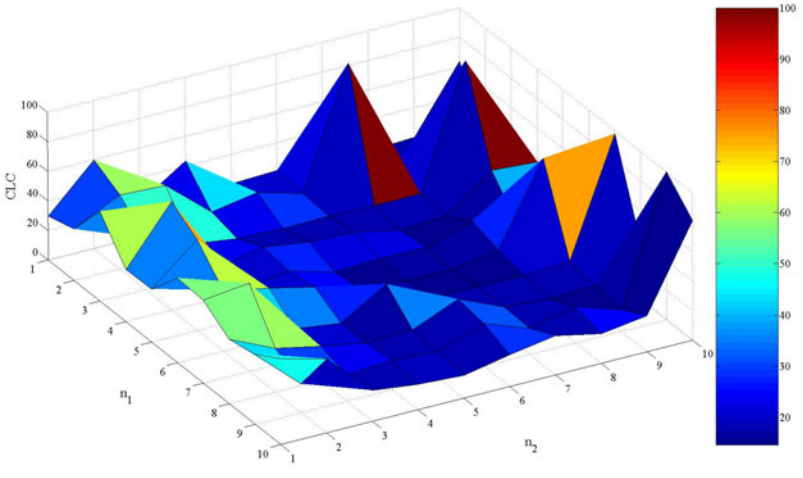


Fig. 6. CLC for PIs constructed using the delta technique (average results over a grid of 100 NNs)

PICP has always been bigger or at least equal to its nominal value (90%). Therefore, all constructed PIs are acceptable in term of their PICP. The shortest PIs are obtained using the delta technique (14.66%, more than 2% less than shortest PIs constructed using the Bayesian technique). In contrast, constructed PIs using the Bayesian technique are in average shorter with smaller standard division. These behaviours can be observed in fluctuations of CLC on the grid of NNs. While for PIs of the Bayesian technique, fluctuations are smooth and decreasing, they are pretty big and erratic for

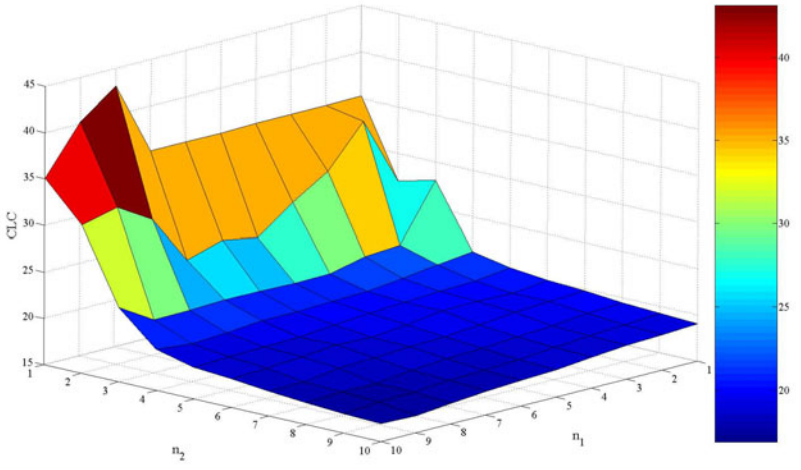


Fig. 7. CLC for PIs constructed using the Bayesian technique (average results over a grid of 100 NNs)

PIs of the delta technique. This indicates that PIs constructed using the delta technique are more affected by initial weights of NN models. For both cases and with the exception of some outliers, there is an overall downward trend in CLCs.

In Fig. 8, PIs of test samples constructed using the delta technique have been demonstrated. The developed NN has 8 and 5 neurons in its hidden layers. CLC, PICP, and NMPIL for this particular case are 18.87, 93.9%, and 18.87%. As PICP is above the nominal confidence level (90%), CLC and NMPIL are the same. NMPIL is less than one fifth of the range of targets. High PICP and small NMPIL means that, (i) PIs well cover the actual values of targets, and (ii), they are satisfactorily narrow and, therefore, can be used instead of unreliable predicted values. One should make note of the fact that upper and lower bounds of PIs can be interpreted and used like the traditional predicted points. The upper bounds of PIs can be interpreted as the highest load demand in the future (the pessimistic case). In contrast, the lower bounds indicate the minimum load required in future (the optimistic case). Based on these interpretations, one may consider both upper and lower bounds when developing power generating schedules. A conservative approach is to consider all upper bounds as future demands and prepare the operation schedule based on them.

It is also interesting to monitor variation of CLC versus network size. In Fig. 9, results shown in Fig. 6 and Fig. 7 have been shown in format of a vector. The top and down plots show CLCs for PIs constructed using the delta and Bayesian techniques, respectively. The horizontal axis is the number of NN parameters. These two plots clearly show that CLCs are highly affected by NN size. As the provided results are averages of five replicates, variations are more likely to be solely due to NN structure not the parameters of NN. The bigger the NNs, the smaller the CLCs, for the Bayesian case. This is due to the promising feature of Bayesian training algorithm. It guarantees that no matter what the size and structure of NNs is,

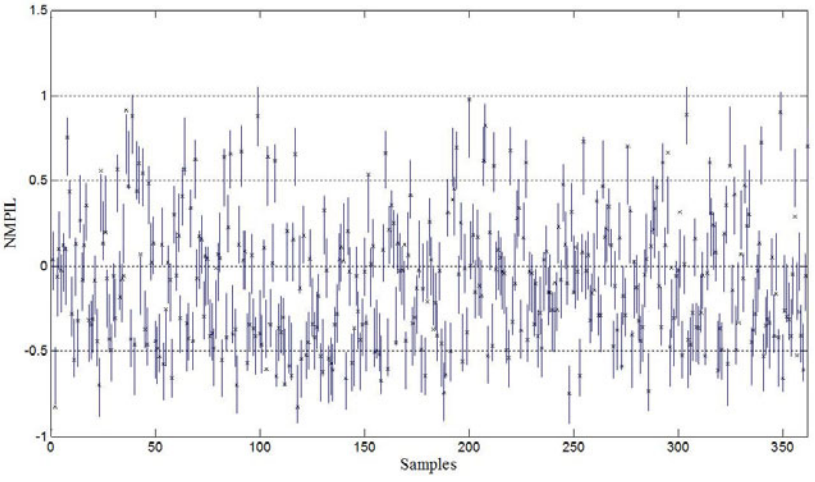


Fig. 8. PIs constructed using the delta technique (NN with (8-5) structure)

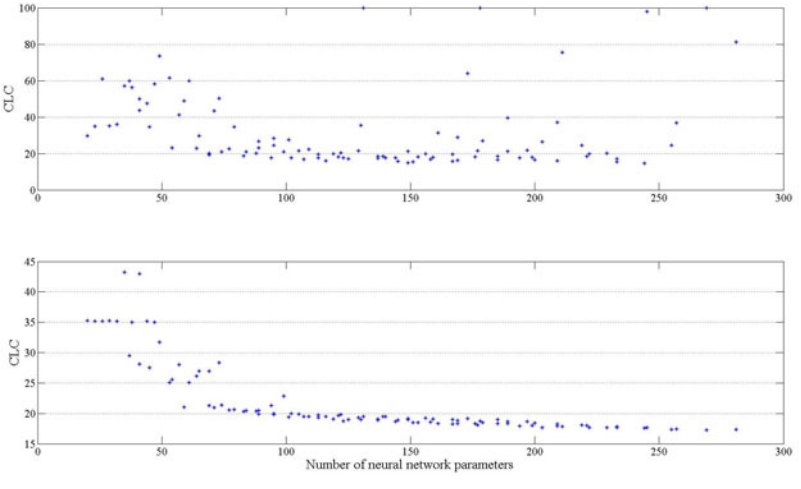


Fig. 9. Distribution of CLC versus NN number of parameters, (top) CLC for PIs constructed using delta technique, (down) CLC for PIs constructed using the Bayesian technique

over-fitting never happens. Therefore, even if the NN size increases, there is no change in the NN performance. From PI construction perspective, this happens for all NNs with more than 100 parameters, and their CLCs are almost the same. For the case of using the delta technique for constructing PIs, the same relationship between CLCs and NN size is observable. Off course, there are cases that they do not follow this principle. Such cases can be considered as highly over-fitted NNs. This means that

even if terms related to the training set in (10) remain small, the generalization power of NNs is highly weakened. Consequently, the PIs are widened.

In our analysis and modeling procedure here, many networks with different number of neurons were investigated. It is a practical aspect of the delta method based on (10) that there is no need to discard any network due to singularity or near singularity of $(J^T J)$ matrix. Our experience shows that in the case of using (6) for constructing PIs, many networks are required to be discarded due to the singularity problem of $(J^T J)$ matrix. This comfort has been achieved through considering λ parameter in (10).

The Bayesian technique for constructing PIs is robust against increase of NN parameters. Still PIs can be optimized to have shorter length (smaller NMPIL) with higher coverage probability. The current common practice is to train NNs to minimize the prediction error and then used the developed NNs for constructing PIs. Instead of this, it is more appropriate to adjust NN parameters on the basis of minimization of CLCs. In this way, PIs are squeezed while guaranteeing their high coverage probability. Such study is the subject of ongoing research.

5 Conclusion

In this study, the problem of load forecasting in power system was studied from a new perspective. Instead of using neural networks for point prediction, NNs were employed for constructing prediction intervals. As these prediction intervals carry more useful information than predicted values of loads, they are more reliable (due to their accuracy indication) and more practical (providing a range of probable loads). The delta and Bayesian techniques were implemented for constructing prediction intervals. Also, a new measure was developed for objective assessment of prediction intervals. The proposed measure simultaneously covers both length and coverage probability of prediction intervals. In the experiments with real data, it was shown that the reliability of neural networks for load forecasting is not very high. At the same, constructed prediction intervals are satisfactorily in term of the proposed measure. Exploring different structures of neural networks revealed that the delta technique yields narrower prediction intervals than the Bayesian technique. In contrast, the Bayesian technique is more robust to the change of neural network structure.

It is our hope that this study will enlighten and inspire researchers to greater empirical study of prediction intervals for modeling and analysis of power systems. Further study may be conducted in different streams. In the area of power engineering, research can be conducted to investigate practicality and reliability of prediction interval for operation planning in power systems. Also, the prediction intervals construction methods and the proposed measure can be applied to similar problems, including price forecasting. From the computational intelligence perspective, it is very important to develop and train neural networks on the basis of minimization of cost functions directly related to the prediction interval qualities rather than point prediction error. Research is ongoing in these directions.

Acknowledgment

This research was fully supported by the Centre for Intelligent Systems Research (CISR) at Deakin University.

References

1. Hobbs, B.F., Helman, U., Jitprapaikulsarn, S., Konda, S., Maratukulam, D.: Artificial neural networks for short-term energy forecasting: Accuracy and economic value. Neurocomputing 23(1-3), 71–84 (1998)
2. Metaxiotis, K., Kagiannas, A., Askounis, D., Psarras, J.: Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher. Energy Conversion and Management 44(9), 1525–1534 (2003)
3. Al-Hamadi, H.M., Soliman, S.A.: Short-term electric load forecasting based on kalman filtering algorithm with moving window weather and load model. Electric Power Systems Research 68(1), 47–59 (2004)
4. Papalexopoulos, A., Hesterberg, T.: A regression-based approach to short-term system load forecasting. IEEE Transactions on Power Systems 5(4), 1535–1547 (1990)
5. Taylor, J.W., de Menezes, L.M., McSharry, P.E.: A comparison of univariate methods for forecasting electricity demand up to a day ahead. International Journal of Forecasting 22(1), 1–16 (2006)
6. Paliwal, M., Kumar, U.A.: Neural networks and statistical techniques: A review of applications. Expert Systems with Applications 36(1), 2–17 (2009)
7. Hornik, K., Stinchcombe, M., White, H.: Multilayer feedforward networks are universal approximators. Neural Networks 2(5), 359–366 (1989)
8. Hussain, M.A.: Review of the applications of neural networks in chemical process control - simulation and online implementation. Artificial Intelligence in Engineering 13, 55–68 (1999)
9. De Gooijer, J.G., Hyndman, R.J.: 25 years of time series forecasting. International Journal of Forecasting 22(3), 443–473 (2006)
10. Bose, B.K.: Neural network applications in power electronics and motor drives—an introduction and perspective. IEEE Transactions on Industrial Electronics 54(1), 14–33 (2007)
11. Lane, V.R., Scott, S.G.: The neural network model of organizational identification. Organizational Behavior and Human Decision Processes 104(2), 175–192 (2007)
12. Bakirtzis, A., Petridis, V., Kiartzis, S., Alexiadis, M., Maissis, A.: A neural network short term load forecasting model for the greek power system. IEEE Transactions on Power Systems 11(2), 858–863 (1996)
13. Chow, T., Leung, C.: Neural network based short-term load forecasting using weather compensation. IEEE Transactions on Power Systems 11(4), 1736–1742 (1996)
14. Mandal, P., Senjyu, T., Urasaki, N., Funabashi, T., Srivastava, A.: A novel approach to forecast electricity price for pjm using neural network and similar days method. IEEE Transactions on Power Systems 22(4), 2058–2065 (2007)
15. Fan, S., Chen, L., Lee, W.-J.: Short-term load forecasting using comprehensive combination based on multimeteorological information. IEEE Transactions on Industry Applications 45(4), 1460–1466 (2009)

16. Barzamini, R., Menhaj, M., Khosravi, A., Kamalvand, S.: Short term load forecasting for iran national power system and its regions using multi layer perceptron and fuzzy inference systems. In: Proceedings of IEEE International Joint Conference on Neural Networks, IJCNN 2005, vol. 4, pp. 2619–2624 (2005)
17. de Aquino, R., Silva, G., Lira, M., Ferreira, A., Carvalho, M., Neto, O., de Oliveira, J.: Combined artificial neural network and adaptive neuro-fuzzy inference system for improving a short-term electric load forecasting. In: de Sá, J.M., Alexandre, L.A., Duch, W., Mandic, D.P. (eds.) ICANN 2007. LNCS, vol. 4669, pp. 779–788. Springer, Heidelberg (2007)
18. Kilmer, R.A., Smith, A.E., Shuman, L.J.: Computing confidence intervals for stochastic simulation using neural network metamodels. Computers & Industrial Engineering 36(2), 391–407 (1999)
19. Zobel, C.W., Keeling, K.B.: Neural network-based simulation metamodels for predicting probability distributions. Computers & Industrial Engineering 54(4), 879–888 (2008)
20. Hwang, J.T.G., Ding, A.A.: Prediction intervals for artificial neural networks. Journal of the American Statistical Association 92(438), 748–757 (1997)
21. de Veaux, R.D., Schumi, J., Schweinsberg, J., Ungar, L.H.: Prediction intervals for neural networks via nonlinear regression. Technometrics 40(4), 273–282 (1998)
22. Wehrens, R., Putter, H., Buydens, L.M.C.: The bootstrap: a tutorial. Chemometrics and Intelligent Laboratory Systems 54(1), 35–52 (2000)
23. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford University Press, Oxford (1995)
24. da Silva, A., Moulin, L.: Confidence intervals for neural network based short-term load forecasting. IEEE Transactions on Power Systems 15(4), 1191–1196 (2000)
25. Papadopoulos, G., Edwards, P., Murray, A.: Confidence estimation methods for neural networks: a practical comparison. IEEE Transactions on Neural Networks 12(6), 1278–1287 (2001)
26. Tibshirani, R.: A comparison of some error estimates for neural network models. Neural Computation 8, 152–163 (1996)
27. Lu, T., Viljanen, M.: Prediction of indoor temperature and relative humidity using neural network models: model comparison. Neural Computing & Applications 18(4), 345–357 (2009)
28. Yu, G., Qiu, H., Djurdjanovic, D., Lee, J.: Feature signature prediction of a boring process using neural network modeling with confidence bounds. The International Journal of Advanced Manufacturing Technology 30(7), 614–621 (2006)
29. Jia, Y., Culver, T.B.: Bootstrapped artificial neural networks for synthetic flow generation with a small data sample. Journal of Hydrology 331(3-4), 580–590 (2006)
30. Ho, S., Xie, M., Tang, L., Xu, K., Goh, T.: Neural network modeling with confidence bounds: a case study on the solder paste deposition process. IEEE Transactions on Electronics Packaging Manufacturing 24(4), 323–332 (2001)
31. Khosravi, A., Nahavandi, S., Creighton, D.: A prediction interval-based approach to determine optimal structures of neural network metamodels. Expert Systems with Applications 37, 2377–2387 (2010)
32. Khosravi, A., Nahavandi, S., Creighton, D.: Constructing prediction intervals for neural network metamodels of complex systems. In: International Joint Conference on Neural Networks (IJCNN), pp. 1576–1582 (2009)
33. Alonso, A.M., Sipols, A.E.: A time series bootstrap procedure for interpolation intervals. Computational Statistics & Data Analysis 52(4), 1792–1805 (2008)
34. Chiu, C.-C., Kao, L.-J., Cook, D.F.: Combining a neural network with a rule-based expert system approach for short-term power load forecasting in taiwan. Expert Systems with Applications 13(4), 299–305 (1997)

35. Zhao, J.H., Dong, Z.Y., Xu, Z., Wong, K.P.: A statistical approach for interval forecasting of the electricity price. IEEE Transactions on Power Systems 23(2), 267–276 (2008)
36. Heskes, T.: Practical confidence and prediction intervals. In: Mozer, T.P.M., Jordan, M. (eds.) Neural Information Processing Systems, vol. 9, pp. 176–182. MIT Press, Cambridge (1997)
37. MacKay, D.J.C.: The evidence framework applied to classification networks. Neural Computation 4(5), 720–736 (1992)
38. Hagan, M., Menhaj, M.: Training feedforward networks with the marquardt algorithm. IEEE Transactions on Neural Networks 5(6), 989–993 (1994)
39. Fay, D., Ringwood, J.V., Condon, M., Kelly, M.: 24-h electrical load data–a sequential or partitioned time series? Neurocomputing 55(3-4), 469–498 (2003)
40. Ghiassi, M., Zimbra, D.K., Saidane, H.: Medium term system load forecasting with a dynamic artificial neural network model. Electric Power Systems Research 76(5), 302–316 (2006)

Neural Network Ensemble for 24-Hour Load Pattern Prediction in Power System

Krzysztof Siwek and Stanislaw Osowski

Abstract. The paper is concerned with the problem of accurate prediction of the 24-hour ahead load pattern in the power system. We propose the solution relying on the application of many different neural predictors combined in an ensemble. Each neural network is responsible for the same job – predicting the 24-hour load pattern for the next day. The series containing 24 values of the load pattern forecasted by each predictor are combined together using principal component analysis, which extracts the most important information and reduces the size of vector used in the final stage of prediction. The final predictor has the form of another neural network. The developed system of prediction was tested on the real data of the Polish Power System. The results have been compared to the appropriate values generated by other methods.

1 Introduction

The prediction of pattern corresponding to 24 hours ahead load demand is very important in industrial research and practice. It is an essential instrument in the reliable and economical generation and distribution of electrical energy, since enables to deliver the really needed power and in this way to reduce the cost of energy. Depending on the time horizon the forecasting can be generally divided into short-term, mid-term and long-term. Short-term load forecasting, ranging from an hour to a week, fulfills important function used in unit commitment, power transfer scheduling and economic emission dispatch. On the other side the mid/long term forecasting covers from a few weeks to several years. It is used for the maintenance, purchasing fuel as well as planning the development of the power plants.

In this paper we will consider the short term load forecasting aiming in prediction of the 24-hour load pattern of the power system. A wide variety of forecasting models suited for this task have been proposed in the past. Most of them can be classified into two broad categories: the statistical methods and artificial intelligence.

Krzysztof Siwek
Warsaw University of Technology – Dept. of Electrical Engineering
ul. Koszykowa 75, 00-661 Warsaw, Poland, Tel. +48222347235

Stanislaw Osowski
Military University of Technology – Dept. of Electronics
ul. Kaliskiego 2, 00-908 Warsaw, Poland
E-mail: {ksiwek,sto}@iem.pw.edu.pl
and
Warsaw University of Technology – Dept. of Electrical Engineering

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 151 – 169.
springerlink.com

To the most known approaches of the first group belong the regression models (ARX, ARMAX, ARIMA), where the load is a random variable function of a set of independent deterministic variables or time series (stochastic or Box-Jenkins models). These approaches consider the load as the result of an stochastic process [10, 28]. The second group is of most interest now and is represented by the neural networks and SVM [2, 12, 13, 18, 27], fuzzy systems [19, 24], evolutionary algorithms [23], expert systems [17] and hybrid models [1, 6, 7, 11, 16, 22] which take advantage of several models.

Nowadays most often used are neural networks due to their capability to approximate any continuous nonlinear function and their self-adaptation properties. The term neural networks has been extended now to many different kinds of neural and neural alike methods, like Multilayer Perceptron (MLP), self-organizing (SO) Kohonen network, Support Vector Machine (SVM), Radial Basis Function (RBF), etc. [9]. Although many different approaches have been developed in the past there is still need to improve the accuracy of the predicting systems.

The neural based methods differ by the particularity of the network structure and the basic nonlinear functions used in prediction, principle of operation and the way of learning. They depend their prediction ability on different aspects of processing of the learning data. For example application of Kohonen network exploits the concept of clusterization of similar data points [4, 15] while the MLP, RBF, Elman networks and SVM take into account the universal approximation ability of these networks [9]. On the other hand MLP network performs the global approximation, the RBF network is a typical local approximation tool, while SVM may apply both, depending on the applied kernel (the sigmoidal, polynomial or Gaussian kernels). The Elman network is a recurrent structure differing significantly from the feedforward ones, like MLP, RBF or SVM. Moreover we should note also great differences in learning algorithms of these types of networks.

The most often used approach in any application of neural networks is to train different networks and then take this one, which guarantees the best reproduction results on the data not taking part in learning (the verification data). Much wiser approach is to take into account all partial prediction results, combine them into one ensemble system of presumably better quality and treat the combined output as the final forecast [14, 21].

This paper develops and investigates the latter philosophy. Instead of discarding the less fortunate prediction results we analyze them and take the conclusions of such analysis into account at preparation of the final forecasting. The ensemble of neural predictors is composed of few individual neural networks. The prediction data generated by each predictor of ensemble are combined together to form common forecasted power pattern for 24 hours ahead. Integration of the individual predictors is achieved here by applying the principal component analysis (PCA) of data and using the most important components as the input signals for the final neural predictor.

This 2-step approach to the power demand forecasting was verified by the numerical experiments concerning the Polish Power System (PPS). The results will be compared to the other methods of forecasting. It will be shown that our approach

leads to the great improvement of the accuracy of prediction. For the same data we were able to reduce the total mean squared errors in a significant way.

2 Analysis of the 24-Hour Load Pattern Prediction Problem

The accurate prediction of the time series corresponding to the electric power consumption in the country belongs to the challenging tasks. It is due to the facts that the load of the power system depends on many different, not controllable factors of statistical nature resulting from the sum of many individual consumers, season of the year, the potentially abrupt changes of weather as well as on some unpredictable events having great impact of temporary consumption of electricity. If we take into consideration the Polish Power System of the last two years we have noted that at the mean load of about P_{mean}=16000MW the observed standard deviation was equal σ=2800MW. Fig. 1 presents the typical hourly change of power consumption of Poland in the last two years.

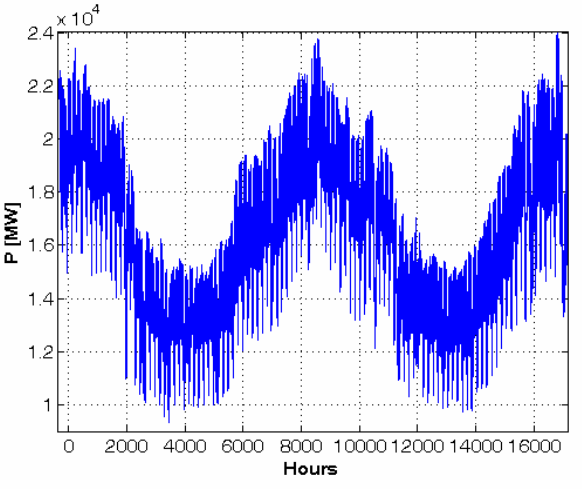


Fig. 1. The change of hourly power consumption of Poland of two years

We can observe the off-peak periods corresponding to late spring and summer months and also peak periods corresponding to autumn and winter months (the heating seasons). In the week scale we can also observe some loads trends: large consumption of power in working days and much smaller need for power at weekends, when most factories stop their activities.

The difficulty of hourly load prediction is well illustrated on the plane of power formed by two consecutive hours. Fig. 2 presents such plot for PPS corresponding to the data of Fig. 1. The horizontal axis represents the power consumption at *h*th hour and the vertical one at (*h*+1)th hour. We can see wide distribution of such points, differing significantly from the diagonal shape.

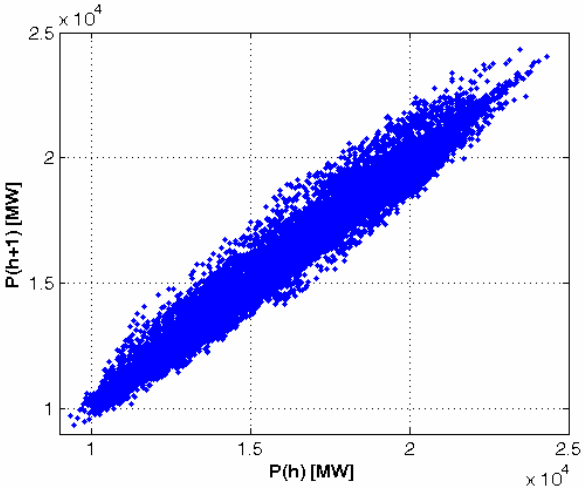


Fig. 2. The distribution of the power consumption in PPS for two consecutive hours

3 The Proposed Neural Approach to Load Prediction

The key point in our approach is using many independent neural predictors. All of them are fed by the same learning data. Their output signals are grouped together and are subject to PCA analysis. Few most important principal components form the input data to the final neural predictor, which delivers the really forecasted values of the load pattern for the next 24 hours.

3.1 The Individual Neural Predictors

In the first step of our solution we apply the ensemble of individual neural predictors learned on the same data set. To get the best results we have to chose predictors of similar quality, which are independent from each other. The independence of their operation is a key point for getting high quality of performance of the ensemble. At independent operation of predictors each of them commit the prediction errors at different points. Taking into account all results at any hour we are able to compensate for same errors and in this way to reduce the total error of prognosis. In our solution we have chosen 4 neural type predictors: the multilayer perceptron, Elman recurrent network, Support Vector Machine and self-organizing Kohonen network. All of them are known from excellent operation in nonlinear signal processing.

3.1.1 The Supervised Neural Predictors

The first three predictors belong to the supervised networks and perform the role of the universal approximators. To represent the generally unknown, next day 24-hour load pattern, they map the past loads of the system into the present forecasted load at dth day and hth hour. The general supervised model of the load forecast has been assumed in the following mathematical form [18]

$$\hat{P}(d,h) = f\left(\mathbf{w}, t, s, P(d,h-1),...,P(d,h-H),P(d-1,h),...,P(d-D,h-H)\right) \quad (1)$$

where **w** represents the vector of parameters of the network, H and D - the number of past hours and days, respectively, influencing the prediction process, t - type of the day (workdays, weekends and holidays) and s - the season of the year (autumn, winter, spring or summer). The symbol $\hat{P}(d,h)$ represents the predicted load and the $P(d-i,h-j)$ written without hat – the known values of the load of the appropriate type of the day and hour from the past. In this mathematical model we have omitted the temperature because the prediction is concerned with the data corresponding to the territory of the whole country. In the country like Poland the temperature is changing a lot in different regions of the country, so it would be difficult to adjust the proper value of the temperature for the particular day. However we should observe that the temperature information is indirectly contained in the power consumption. On the other side in the case of forecasting the energy consumption for a small region of the country the inclusion of temperature and gradient of temperature in the model would be beneficial and easy to consider in our approach.

To provide the similar impact of all input variables the data samples should be normalized. The normalization may take different forms, from which the simplest one is to divide the real load by the mean value of the data base of the Power System, corresponding to the years taking part in experiments.

The particular form of the applied predictors depends on their structure and way of learning. The expression (1) may be associated with MLP network of the particular structure presented in Fig. 3. The MLP [9] consists of many simple neuron-like processing units of sigmoidal activation function grouped together in layers. The synaptic connections exist only between the neighboring layers. The number of hidden layers and neurons of sigmoidal non-linearity are usually subject to adjustment in an experimental way by training different structures and choosing the smallest one, still satisfying the learning accuracy. The output layer contains 24 linear neurons. Their number is equal to the number of hours of prediction (24 hours ahead). The signals of output neurons represent the normalized forecasted 24-hour load pattern. The learning procedure of MLP is reduced to the minimization of the Euclidean error measure between the actual and desired output signals over the entire learning data set. The most effective learning approach applies the gradient information and uses the first or second order optimization algorithms, like Levenberg-Marquard or conjugate gradient [9]. Gradient vector in multilayer network is computed using the backpropagation algorithm.

The input layer represents the excitation nodes. One node is used for binary coding of the type of day (working day or non-working day) and two - the season of the year (winter, spring, summer and autumn). Most nodes represent the loads of some past days (up to D) and previous hours (up to H). All these signals put to the input nodes form the input vector **x**.

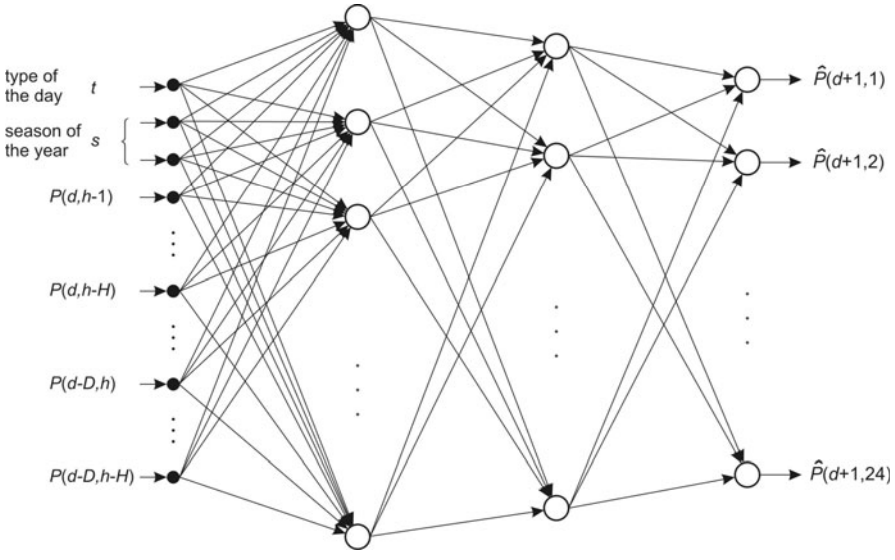


Fig. 3. The multilayer perceptron structure used at forecasting the 24 hour load pattern

The important point in this approach is adjusting the length of the input vector **x**, which depends explicitly on the chosen values of D and H. There are some methods for optimal selection of the input variables, based on data mining [2, 8]. To such tools belong: the correlation analysis, PCA, projection pursuit, application of the linear SVM feature ranking, etc [8]. In this work we have applied the correlation analysis studying the degree of correlation of the learning errors of MLP with different numbers of D and H. On the basis of the numerical simulations we have found that optimal number of input nodes in our case is 18, which corresponds to D=3 and H=3.

The optimal number of hidden layers and neurons in these layers was found by using trial and error approach. It means learning many different structure MLP networks and accepting this one which provides the least value of the error on the validation data extracted from the learning data set (usually 20% of the learning set). On the basis of these experiments we have found the optimal structure consisting of two hidden layers of 20 and 15 sigmoidal neurons, respectively. In this way the optimal structure of MLP network used in prediction is described as 18-20-15-24.

Elman network is a two layer recurrent structure of sigmoidal neurons, implementing the feedback from the hidden layer to the input of the network [9]. This feedback path allows Elman network to learn to recognize and generate temporal patterns, which are of interest in prediction. The external input signals to Elman network are the same as for the MLP network. The actually applied network structure is shown in Fig. 4.

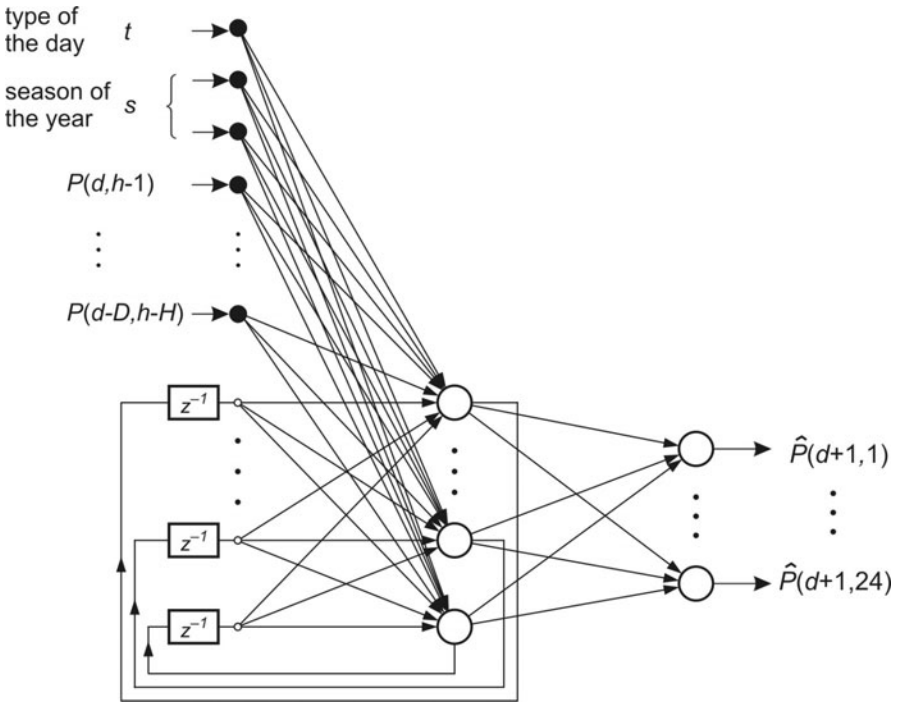


Fig. 4. The Elman structure used for forecasting the 24-hour load pattern

The most distinct difference of Elman network to the MLP is the feedback. Thanks to its existence we are able to reduce the complexity of the network. In the numerical experiments we have used the structure 18-8-24 applying 18 input external nodes containing the same signals as in MLP, only 8 hidden neurons and 24 output neurons. The significant reduction of the synaptic weights was possible thanks to the feedback. The learning strategy of Elman network uses similar principle of minimization of error function defined for the learning data, implemented in the same way as for MLP. In our experiments it was performed using the Levenberg-Marquardt algorithm [9].

Support Vector Machine is another powerful neural like structure developed by Vapnik [25]. We have used it in a regression mode. Note that SVM contains only one output unit, hence to solve the prediction of the load pattern for 24 hours ahead we have to train 24 separate SVM networks. In training them we use the same input data arranged in an 18-element vectors **x**, identically as for MLP and Elman networks. In distinction to the previous networks SVM does not predefines the number of hidden units (kernel functions). This number is automatically determined in the learning procedure.

The learning strategy of SVM network is relied on another philosophy than in the MLP and Elman networks. Instead of minimizing the error function defined for the learning data it minimizes the weights of the network, while keeping the output signals as close as possible to their destination values [20, 25]. Application

of Lagrange function to such problem leads to the final quadratic optimization task (so called dual problem), involving Lagrange multipliers α_i, α_i' (i=1, 2, …, p) responsible for the functional constraints

$$\max \sum_{i=1}^{p} t_i(\alpha_i - \alpha_i') - \varepsilon \sum_{i=1}^{p}(\alpha_i + \alpha_i') - \frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}(\alpha_i - \alpha_i')(\alpha_j - \alpha_j')K(\mathbf{x}_i, \mathbf{x}_j) \tag{2}$$

at the constraints

$$\begin{aligned} &\sum_{i=1}^{p}(\alpha_i - \alpha_i') = 0 \\ &0 \le \alpha_i \le C \quad 0 \le \alpha_i' \le C \end{aligned} \tag{3}$$

where p is the number of learning data ($\mathbf{x}_i$, t_i), $K(\mathbf{x}_i, \mathbf{x}_j) = \varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}_j)$ the inner-product kernel function defined in accordance with Mercer's theorem [20], $\boldsymbol{\varphi}(\mathbf{x})$ the activation function vector, ε the assumed tolerance and C the user specified regularization parameter. The variables ε and C are free hyperparameters that control the Vapnik-Chervonenkis dimension of the approximating function and influence the generalization abilities of SVM. Both must be selected a'priori by the user.

The solution of the dual (quadratic) problem with respect to Lagrange multipliers is relatively easy and there are many very efficient algorithms leading to the global minimum [20]. After solving the dual problem we can determine the network output signal $y(\mathbf{x})$ through the optimized Lagrange multipliers α_i, α_i' and the kernel function $K(\mathbf{x},\mathbf{x}_i)$ in the following way [20]

$$y(\mathbf{x}) = \sum_{i=1}^{N_{sv}}(\alpha_i - \alpha'_i)K(\mathbf{x}, \mathbf{x}_i) + b \tag{4}$$

where N_{sv} is the number of so called support vectors (the number of nonzero Lagrange multipliers) and b is the bias. In practice the most often used kernels are the Gaussian and polynomial functions.

To get reliable results of learning we have to make proper choice of hyperparameters ε and C. Constant ε determines the margin within which the error is neglected. The smaller its value the higher is the accuracy of the required matching of the response $y(\mathbf{x}_i)$ of SVM to the proper target values t_i in learning data set. However, too accurate matching of the learning data may result in the loss of generalization ability of the network, leading to the increase of the testing error. The constant C is the weight, determining the balance between the complexity of the network characterized by the weight vector **w,** and the error of approximation. For the normalized input signals the value of ε is usually adjusted in the range ($10^{-3} - 10^{-2}$) while C is much bigger than the value of 1 (the typical value 1000).

3.1.2 The Self-organizing Approach to Prediction

To differentiate the types of predictors as much as possible we have additionally applied the self-organizing Kohonen network approach [4, 9]. This network is

learned by using competitive algorithms [9]. The main task of the self-organizing network is to learn the characteristics of the daily loads (profile vectors) of the system, defined in the way [4]

$$p(d,h) = \frac{P(d,h) - P_m(d)}{\sigma(d)} \tag{5}$$

for each day d and 24 hours h=1, 2, ..., 24. The symbol $P(d,h)$ is the real load of dth day at hth hour, $P_m(d)$ is the mean value of the load of dth day and $\sigma(d)$ the standard deviation of the load of dth day. The set of 24-load profiles forms the profile vector for dth day $\hat{\mathbf{p}}(d) = \left[\hat{p}(d,1), \hat{p}(d,2),..., \hat{p}(d,24)\right]^T$.

The most important task of Kohonen network is to group similar data vectors into clusters. The days of the same type belonging to the same seasons of the year have similar profile patterns and are grouped together. Each cluster is represented by one neuron, acting in the competitive mode. Once the network is trained, each neuron represents the data closest to its weight vector in the chosen metric space. The prediction of the load for dth day and hth hour may be expressed now in the form

$$\hat{P}(d,h) = \hat{\sigma}(d)\hat{p}(d,h) + \hat{P}_m(d) \tag{6}$$

where the variables with hat mean the predicted values. To make the prediction of the load $\hat{P}(d,h)$ for the particular day and hour we have to know not only the load profile $\hat{p}(d,h)$, but also the mean value $P_m(d)$ and standard deviation $\hat{\sigma}(d)$ of the load for this day.

The predictions of the mean value and standard deviation for the particular day have been obtained by applying the standard MLP network [18] in a way very similar to the already presented 24-hour load pattern approach. Fig. 5 presents the general structure of this network used for prediction of the mean value of the dth

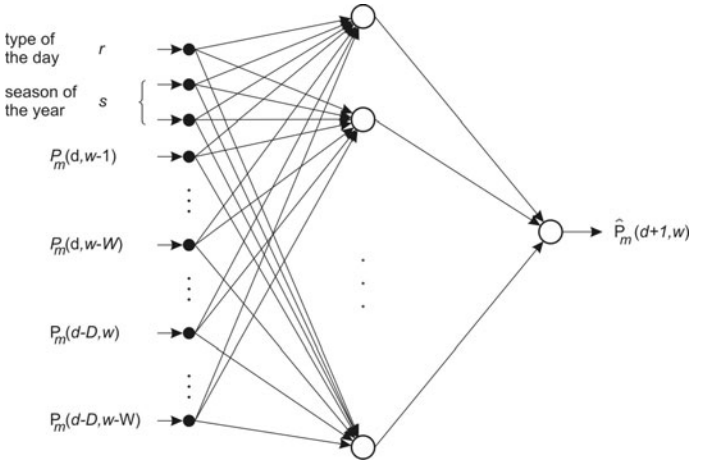


Fig. 5. The MLP network structure used for prediction of the daily mean load

day and of *w*th week. The mean daily loads taken into account correspond to *D* past days of the same week and *D*+1 days of the previous *W* weeks.
Similar structure is used for standard deviation prediction. The only difference is substitution of the mean values by the standard deviations of the daily load.
The profile vector prediction for *d*th day is estimated by averaging the winner vectors of the Kohonen network for this particular day (for example Tuesdays in July) from the past history, i.e.,

$$\hat{\mathbf{p}}(d) = \frac{\sum_{i=1}^{n} k_{di}\mathbf{w}_i}{\sum_{i=1}^{n} k_{di}} \tag{7}$$

where k_{di} is the quantity of appearances of *i*th neuron among the winners in the past for this particular day type and $\mathbf{w}_i$ is the weight vector of the winner.

3.2 Principal Component Analysis of Data

Each neural predictor described in the previous section generates the series of 24 load values predicted for 24 hours of the next day. Our task in the next step is to compress the amount of this data by keeping the essential part of the information. We perform this step by applying the principal component analysis of the data [5].

Consider a data set generated by *M* predictors arranged in the form of the matrix **Z** of the size $p \times (24M)$, where *p* is the number of days under prediction and *M* the number of predictors (in our case *M*=4). Each predictor output for one day contains 24 predicted power demands for the particular hours of this day. The rows of the matrix **Z** are the composition of the concatenated outputs of *M* predictors for the respective day. The aim of the principal component analysis is to map these *p* high-dimensional vectors into a lower dimensional space [5, 9]. In this way each long vector **z** (the rows of the matrix **Z**) will be represented now by the vector **y** of smaller dimension *K*, containing sufficiently high percentage of the most important part of the original information.

On the basis of the learning data set we form first the auto covariance matrix $\mathbf{R}_{zz} = \frac{1}{p}\sum_{k=1}^{p} \mathbf{z}_k \mathbf{z}_k^T$, then perform the eigen-value analysis of this matrix

$$\mathbf{R}_{zz} = \sum_{k=1}^{N} \lambda_i \mathbf{w}_i \mathbf{w}_i^T \tag{8}$$

and at the define the PCA matrix $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_K]^T$ formed from *K* most important eigen vectors $\mathbf{w}_i$ (*i*= 1, 2, …, *K*) associated with the largest eigenvalues. The transformation of 24*M*-dimensional vectors **z** into lower *K*-dimensional vectors **y** is defined by a simple linear relation

$$\mathbf{y} = \mathbf{W}\mathbf{z} \tag{9}$$

The transformation defined in this way determines the low-dimensional vector **y**, representing the essential part of information concerning the load pattern for the next 24 hours.

3.3 Final Predictor

The set of p low-dimensional vectors $\mathbf{y}$ is used in the next stage as the input training data for the final predictor, whose output signals will represent the finally forecasted 24-component time series of the next 24 hours of the day. To get high quality of prediction results we have to use the predictor of highest possible accuracy. On the basis of results of the first stage prediction we decided to check two best neural structures: the MLP and SVM. In the case of MLP only one network of 24 linear output neurons is used. Applying SVM we have to train 24 SVM structures (each specializing for the particular hour of the day) fed by the same input data. As the learning data for the final predictor we have used the pairs $(\mathbf{y}_i, \mathbf{t}_i)$ for i=1, 2, ..., p. Vectors $\mathbf{y}_i$ result from PCA analysis and $\mathbf{t}_i$ are the known load patterns used also in learning the neural predictors in the first step of our approach. Fig. 6 presents the final structure of forecasting system described in this paper.

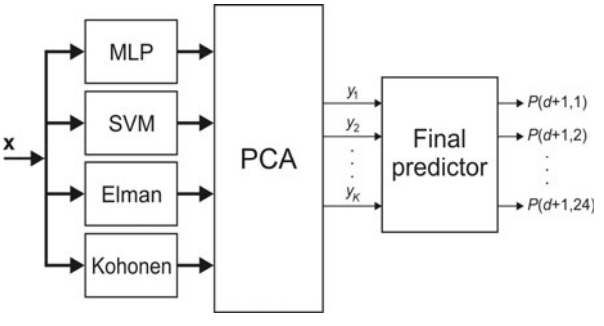


Fig. 6. The diagram of the proposed 2-stage forecasting system

It contains 4 individual predictors (MLP, SVM, Elman and Kohonen networks). The PCA fed by 24M (M=4) signals reduces their size to K signals. These signals form the excitation of the second stage neural predictor, delivering the final forecasted values of the load pattern for 24 hours of the next day.

4 Numerical Results

4.1 The Data Base

The numerical experiments have been performed for the data of the Polish Power System of three years (over 26280 hours). The same data set applied in learning and testing has been used for each individual predictor. The first two years (17520 hours) have been applied only in the learning stage and the last year, not taking part in learning (8760 hours) has been left for testing all trained predictors. The data samples have been normalized dividing the real load by the mean value of the data base of the Polish Power System of 3 years taking part in the experiments.

4.2 The Neural Predictors

The individual predictors have been adapted on the basis of the learning data. The optimal MLP network of the structure 18-20-19-24 was trained using the conjugate gradient algorithm [9]. The input signals of the neural network have been formed by the normalized loads of the nearest past 3 hours of the actual day and 4 hours (the actual hour and 3 nearest past) for 3 previous days of the same type (15 components together), as well as the type of the season (two nodes coded in binary way: 00 – spring, 01 - summer, 10 – autumn an 11 – winter) and type of the day (one node: 1 – working days, 0 – non-working days). The hidden neurons were sigmoidal and the output neurons linear. The particular numbers of hidden neurons have been determined in the introductory stage of experiments using the validation data set extracted from the learning data (1/5 of the set). Each output neuron was responsible for prediction of the load for the particular hour of the day. The network was learned by applying the conjugate gradient algorithm, implemented on Matlab platform [29].

The Elman recurrent network structure (18-8-24) applied also 18 input nodes containing the same signals as in MLP. The structure was composed of 8 hidden neurons and 24 output neurons. The network was learned by applying Levenberg-Marquardt algorithm implemented on Matlab platform [29].

The SVM network of also 18 inputs applies special strategy of prediction since it possesses only one output neuron. We have trained 24 separate SVM networks of Gaussian kernel functions $K(\mathbf{x},\mathbf{x}_i) = e^{-\gamma\|\mathbf{x}-\mathbf{x}_i\|^2}$ for prediction of 24-point time series (each SVM network responsible for prediction of the load of the particular hour of the day). The hyperparameters applied in prediction were: C=1000, γ=0.5 and ε=0.01. The modified Platt algorithm implemented on Matlab platform was applied in learning all SVM networks [20].

In the self-organizing approach we have applied 100 self-organizing neurons for prediction of the profiles. This number was found after series of introductory experiments. After adapting the weights of Kohonen network the learning data of all days have been tested and the winners determined for the profile vectors of all days. These winners are then used for prediction of the profile vector for the particular day in the future (relation (7)). The MLP networks responsible for prediction of the mean values and standard deviations of the load for each day were of the following structures: 10-6-1 (the mean values) and 14-8-1 (standard deviation). The input vectors for both networks have been arranged by applying the same philosophy as in direct MLP prediction. In prediction of the mean we used the daily mean loads of previous 3 days of the same week and of 4 days (the actual and 3 previous days) of the previous week. Additionally we have used 2 nodes to code the season of the year and one node to code the type of the day. The MLP network for forecasting the standard deviation used the same structure of data plus the data of the additional week (two previous weeks instead of one) of the past.

The PCA transformation of the data was tried at different number of principal components. The best results have been obtained at K=24 and this number was applied in the final experiments to form the excitation for the final neural predictor. Two different neural networks have been checked for final prediction: the MLP and SVM. The results of both experiments will be presented and discussed in the next sections.

4.3 The Measures of Prediction Errors

The results of prediction have been compared on the basis of the committed errors. There are four most important (from the practical point of view) types of errors. Let us denote by $P(h)$ and $\widehat{P}(h)$ the real and predicted load at hth hour, respectively, and by n the total number of hours under prediction. We have adopted the following definitions of errors:

- the mean absolute percentage error (MAPE)

$$MAPE = \frac{1}{n}\sum_{h=1}^{n}\frac{\left|P(h)-\widehat{P}(h)\right|}{P(h)}\cdot 100\% \tag{10}$$

- the mean squared error (MSE)

$$MSE = \frac{1}{n}\sum_{h=1}^{n}\left[P(h)-\widehat{P}(h)\right]^2 \tag{11}$$

- the normalized mean squared error (NMSE)

$$NMSE = \frac{MSE}{\left[mean(P)\right]^2} \tag{12}$$

where $mean(P)$ represents the mean value of the load in the time period of prediction

- the maximum percentage error (MAXPE)

$$MAXPE = \max\left\{\frac{\left|P(h)-\widehat{P}(h)\right|}{P(h)}\cdot 100\%\right\} \tag{13}$$

These errors have been calculated separately for the learning and testing data. Here we will limit ourselves only to the testing errors, related to the data not taking part in learning, since this information is the most objective way of assessing the proposed predictive system. Taking into account the stochastic nature of the learning algorithms used in training of the neural networks we have repeated the learning and testing procedures 20 times starting from random values of weights. After all trials we have determined the means and standard deviations of these errors.

4.4 The Numerical Results of Prediction

The results of prediction will be presented for each individual predictor and for the ensemble. Table 1 presents the obtained values of the mean and standard deviation of the testing errors for all individual predictors corresponding to one year (365 days and 8760 hours) not taking part in learning. Note that SVM learning algorithm at constant values of the hyperparameters delivers repeatable results (hence *std* is equal zero).

Table 1. The testing errors of the load forecasting for 8760 hours of the Polish Power System by using individual predictors

Method	MAPE [%]	MAXPE [%]	MSE [MW^2]	NMSE
MLP	2.06±0.13	16.95	(1.74±0.12)e5	(6.81±0.46)e-4
SVM	2.21±0	28.30	(2.92±0)e5	(11.6±0)e-4
Elman	2.24±0.09	24.97	(3.13±0.10)e5	(12.0±0.41)e-4
SO	2.36±0.03	18.12	(2.43±0.15)e5	(9.33±0.50)e-4

It is evident that in any respect the most accurate is the MLP predictor, although the other solutions are not far from it. All MAPE errors are comparable to each other and this provides good premise for their successive integration in the ensemble.

The results generated by all individual predictors have been processed according to the described procedure: first by PCA (24 main principal components selected) and then by second stage prediction using independently MLP o the structure 24-10-24, and SVM as the final predictors. The mean errors and standard deviations of the final prediction results, performed 20 times on the testing data are gathered in Table 2.

Table 2. The testing errors of the final load forecasting by the ensemble of predictors for the Polish Power System obtained in 20 trials

	MAPE [%]	MAXPE [%]	MSE [MW^2]	NMSE
MLP final predictor	1.47±0.09	14.19	(1.042±0.06)e+5	(4.06±0.36)e-4
SVM final predictor	1.34±0.05	10.65	(0.947±0.07)e+5	(3.69±0.21)e-4

The obtained results confirm very high efficiency of the ensemble of predictors. All error measures have been reduced significantly in comparison to the best individual predictor (MLP) The MAPE was reduced from 2.06% (the best individual predictor) to 1.34% (the ensemble employing SVM integration). It means more than 28% relative improvement. Even more spectacular is the relative improvement of MSE and MAXPE error. The MSE has been reduced by 46% and MAXPE by 36% in relation to the best individual predictor.

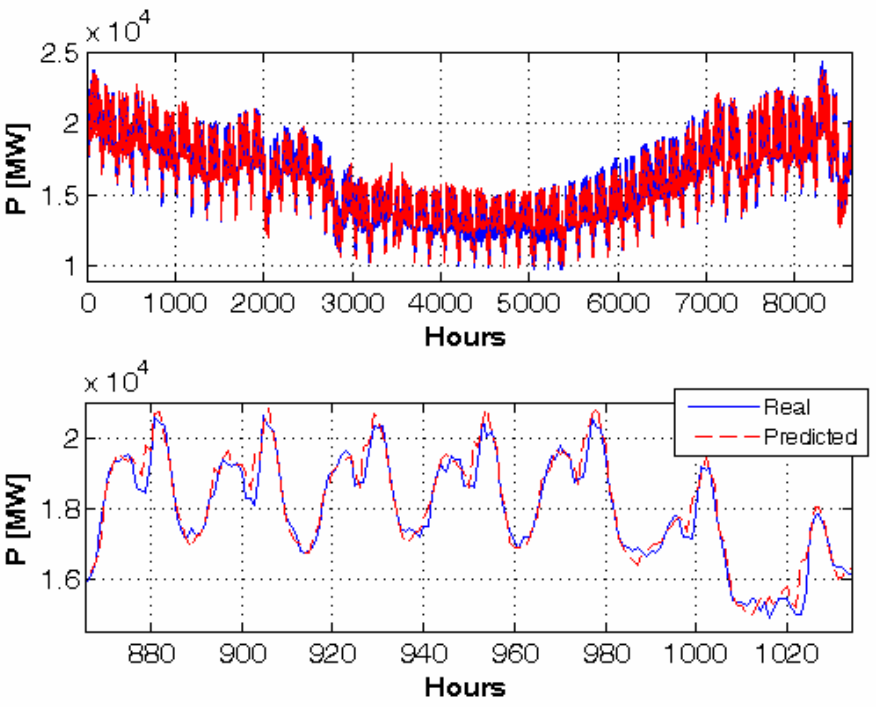


Fig. 7. The graphical presentation of the prediction results related to ensemble: a) the superimposed predicted and real load pattern of the whole year, b) the details of prediction corresponding to one week of the year.

Fig. 7 presents the graphical results of the power demand prediction corresponding to the data of the last year not taking part in learning. They are related to the best ensemble of predictors. The upper figure shows the predicted values compared to the really measured load, and bottom figure presents one chosen week segment just to show the details of both curves.

There are visible load patterns corresponding to 7 days of the week. The level of power consumption for the first 5 days (from Monday to Friday) is significantly higher than that for Saturday and Sunday (the last two segments). The load patterns of the working days are to some degree similar. On the other side the shapes of the patterns for Saturday and Sunday are significantly different.

4.5 Comparison of Results to Other Integration Techniques

The paper has shown the application of many predictors for getting the accurate forecast of the 24-hour load pattern. The natural question arises how the proposed method of integration is related to the other possible solutions. The most natural way of integration of the results of many predictors is the simple averaging (ordinary mean) for every hour under prediction. At M predictors the forecasted value of the load for dth day at hth hour may be then calculated as

$$\hat{P}(d,h) = \frac{1}{M}\sum_{i=1}^{M}\hat{P}_i(d,h) \tag{14}$$

At different accuracy of predictors better results may be expected by applying the weighted averaging, taking into account the accuracy of each predictor. The final forecast can be then described by

$$\hat{P}(d,h) = \sum_{i=1}^{M} w_i \hat{P}_i(d,h) \tag{15}$$

In this expression w_i is the weight associated with ith predictor calculated on the basis of the learning results. In the case of comparable accuracy η of each predictor the most reasonable seems to be the simple formula

$$w_i = \frac{\eta_i}{\sum_{k=1}^{M}\eta_k} \tag{16}$$

Another more complex form of integrating the results of different predictors have been presented in [17]. It applies blind source separation (BSS) of data [3]. The BSS system decomposes the original stream of predicted values of the hourly load produced by each individual predictor into independent components by using linear transformation

$$\mathbf{Y} = \mathbf{W}\mathbf{X} \tag{17}$$

In this expression $\mathbf{X}$ represents the matrix of M rows (M – the number of predictors) and n columns (n – the number of hours under prediction) and $\mathbf{W}$ the BSS matrix, $\mathbf{W} \subset \mathbf{R}^{M\times M}$. Each row of the matrix $\mathbf{Y}$ represents the independent component series. Some of these series contain the essential information and some the noise or redundant information (from the point of view of prognosis). Cutting insignificant components and reconstructing the original time series back into the real prognosis on the basis of the essential independent components only, will provide the prognosis deprived of the noise, i.e., of presumably better quality. The reconstruction $\hat{\mathbf{X}}$ of the original data matrix $\mathbf{X}$ is done by using inverse operation called deflation [3]

$$\hat{\mathbf{X}} = \mathbf{W}^{-1}\hat{\mathbf{Y}} \tag{18}$$

in which $\hat{\mathbf{X}}$ denotes the reconstructed time series matrix and $\hat{\mathbf{Y}}$ - the independent component matrix, formed from the original matrix $\mathbf{Y}$ by zeroing some rows, representing the noise. As a result of reconstruction we get M time series representing the final solutions, since the number of reconstructed variables is the same as the number of inputs. We accept the channel, which represents the best accuracy for the learning data set. Table 3 depicts the comparison of the presented approach to these three techniques by applying exactly the same data set. The superiority of the actually presented way of integration is evident.

Table 3. The comparison of results of different methods of integration of predictors

Method	MAPE [%]	MAXPE [%]	MSE [MW^2]	NMSE
Ordinary mean	1.86	16.98	1.48e5	5.76e-4
Weighted average	1.84	16.97	1.47e5	5.72e-4
BSS	1.73	16.21	1.25e5	4.87e-4
Presented approach	**1.34**	**10.65**	**0.947e+5**	**3.69e-4**

It is interesting to compare also the accuracy of our results to different approaches presented by other authors. The results of prediction for the same data of the Polish Power System have been presented in [22, 26]. These papers have given the results only in the form of the normalized mean squared error (NMSE) defined as the real MSE value divided by the square of the mean value. The best resulting NMSE of [22] was equal NMSE=1.6e-3. In the case of [26] the best result of NMSE was 1.8e-3. Our best result corresponding to the same data was equal 0.369e-3.

5 Conclusions

The paper has presented the novel approach to forecasting the 24-hour load pattern of the power system. In the proposed solution many different predictors are trained simultaneously and their results combined together using principal component analysis and additional neural integrator. In the classical approach the less fortunate predictors are usually discarded and the results of the best one are treated as the final outcome. In the presented approach we analyze all of them and take into account at preparation of the final forecast. In the presented solution we have used four individual predictors, although this number may be easily extended without any significant changes of the general procedure. The best practice is to apply predictors relying their outcome on different principle of signal processing. Thanks to such choice the results of their prediction will be statistically independent.

To get the best possible results of integration of many predictors we have applied two stage procedure. In the first step we apply principal component analysis to reduce the size of data vectors while keeping the most essential part of information. In the second step we employ the additional neural predictor to make the final forecast. The best results in this step have been obtained at application of the SVM integrating system.

The experimental results have shown that the performance of the individual predictors can be improved significantly by proper integration of their outputs. The improvement is observed even at application of different quality predictors. For the data corresponding to the Polish Power System and application of four different predictors we have got 28% relative improvement of MAPE and more than 40% of MSE error over the best individual predictors (the MLP network).

It is worth noting that the same principle of integration may be applied for any type of forecasting problems, for example in financial business, economy, environmental pollution, etc. Thanks to the proposed solution we may expect significant increase of accuracy of predictive system.

Acknowledgments. This work is supported by Polish Ministry of Science and Higher Education by grant in the years 2008-2010.

List of Abbreviations

BSS - Blind Source Separation
MAPE - Mean Absolute Percentage Error (MAPE)
MAXP - Maximum Percentage Error
MLP - Multilayer Perceptron
MSE - Mean Squared Error
NMSE - Normalized Mean Squared Error
PCA – Principal Component Analysis
PPS - Polish Power System
RBF - Radial Basis Function
SO - Self-organizing network.
SVM - Support Vector Machine

References

1. Bashir, Z.A., El-Hawary, M.E.: Applying wavelets to short-term load forecasting using PSO-based neural networks. IEEE Transactions on Power Systems 24, 20–27 (2009)
2. Chen, B.J., Chang, M.W., Lin, C.J.: Load forecasting using support vector Machines: a study on EUNITE competition. IEEE Transactions on Power Systems 19, 1245–1248 (2004)
3. Cichocki, A., Amari, S.I.: Adaptive blind signal and image processing. Wiley, New York (2003)
4. Cottrell, M., Girard, B., Girard, Y., Muller, C., Rousset, P.: Daily electrical power curve: classification and forecasting using a Kohonen map. In: Sandoval, F., Mira, J. (eds.) IWANN 1995. LNCS, vol. 930, pp. 1107–1113. Springer, Heidelberg (1995)
5. Diamantras, K., Kung, S.Y.: Principal component neural networks. Wiley, New York (1996)
6. Fidalgo, J.N., Pecas Lopez, J.: Load forecasting performance enhancement when facing anomalous events. IEEE Trans. Power Systems 20, 408–415 (2005)
7. Gonzalez-Romera, E., Jaramillo-Moran, M.A., Carmona-Fernandez, D.: Monthly electric energy demand forecasting based on trend extraction. IEEE Trans. Power Systems 21, 1946–1953 (2006)
8. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. Journal of Machine Learning Research 3, 1158–1182 (2003)
9. Haykin, S.: Neural networks, a comprehensive foundation. Macmillan, New York (2002)

10. Hippert, H.S., Pedreira, C.E., Souza, R.C.: Neural networks for short-term load forecasting: a review and evaluation. IEEE Trans. on Power Systems 16, 44–55 (2001)
11. Hong, W.C.: Hybrid evolutionary algorithms in a SVR-based electric load forecasting model. International Journal of Electrical Power & Energy Systems 31, 409–417 (2009)
12. Kandil, N., Wamkeue, R., Saad, M., Georges, S.: An efficient approach for short term load forecasting using artificial neural networks. Electrical Power and Energy Systems 28, 525–530 (2006)
13. Khotanzad, A., Rohani, R.A., Maratukulam, D.: ANNSTLF–Artificial Neural Network Short-Term Load Forecaster–generation three. IEEE Transactions on Neural Networks 13, 1413–1422 (1998)
14. Kuntcheva, L.: Combining pattern classifiers - methods and algorithms. Wiley, New Jersey (2004)
15. Mandal, P., Senjyu, T., Urasaki, N., Funabashi, T.: A neural network based several hours ahead electric load forecasting using similar days approach. Electrical Power and Energy Systems 28, 367–373 (2006)
16. Mao, H., Zeng, X.J., Leng, G., Zhai, Y.J., Keane, J.: Short term and midterm load forecasting using a bilevel optimization method. IEEE Transactions on Power Systems 24, 1080–1085 (2009)
17. Osowski, S., Siwek, K., Szupiluk, R.: Ensemble neural network approach for accurate load forecasting in the power system. Applied Mathematics and Computer Science 19, 303–315 (2009)
18. Osowski, S., Siwek, K.: Regularization of neural networks for load forecasting in power system. IEE Proc. GTD 149, 340–345 (2002)
19. Papadakis, S.E., Theocharis, J.B., Bakirtzis, A.G.: A load curve based fuzzy modeling technique for short-term load forecasting. Fuzzy Sets and Systems 135, 279–303 (2003)
20. Schölkopf, B., Smola, A.: Learning with Kernels. MIT Press, Cambridge (2002)
21. Siwek, K., Osowski, S.: Two-stage neural network approach to precise 24-hour load pattern prediction. In: Corchado, E., Wu, X., Oja, E., Herrero, Á., Baruque, B. (eds.) HAIS 2009. LNCS, vol. 5572, pp. 327–336. Springer, Heidelberg (2009)
22. Sorjamaa, A., Hao, J., Reyhani, N., Li, Y., Lendasse, A.: Methodology for long-term prediction of time series. Neurocomputing 70, 2861–2869 (2007)
23. Subbaraj, P., Rajasekaran, V.: Evolutionary techniques based combined artificial neural networks for peak load forecasting. Int. Journal of Electrical Power and Energy Systems Engineering 1, 88–94 (2008)
24. Tamimi, M., Egbert, R.: Short term electric load forecasting via fuzzy neural collaboration. Electric Power Systems Research 56, 243–248 (2000)
25. Vapnik, V.: Statistical learning theory. Wiley, New York (1998)
26. Wertz, V., Lendasse, A., Cottrell, M., Verleysen, M.: Prediction of electric load using Kohonen maps – application to the Polish electricity consumption. In: American Control Conference, Anchorage, pp. 3684–3688 (2002)
27. Yalcinoz, T., Eminoglu, U.: Short term and medium term power distribution load forecasting by neural networks. Energy Conversion and Management 46, 1393–1405 (2005)
28. Yang, H.T., Huang, C.M.: A new short-term load forecasting approach using self-organizing fuzzy ARMAX models. IEEE Transactions on Power Systems 13, 217–225 (1998)
29. Matlab manual, user's guide, MathWorks, Natick, MA (2008)

Power System Protection Using Machine Learning Technique

S. R. Samantaray, P.K. Dash, and G. Panda

Abstract. This chapter presents a new approach for distance relaying of transmission line using machine intelligence technique such as Support Vector Machine (SVM). SVM is a relatively new computational learning method based on the statistical learning theory. The proposed technique is used for developing protection schemes for Thyristor Controlled Series Compensated (TCSC) Line using post fault current samples for half cycle (10 samples) from the inception of the fault and firing angle as inputs to the SVM. Three SVMs are trained to provide fault classification, ground detection and section identification respectively for the TCSC line. Also SVM is used for faulty phase selection and ground detection in large power transmission system without TCSC. The method uses post fault current and voltage samples for $1/4^{th}$ cycle (5 samples) as inputs to SVM-1 to result faulty phase selection. SVM-2 is trained and tested with zero sequence components of fundamental, 3^{rd} and 5^{th} harmonic components of the post fault current signal to result the involvement of ground in the fault process. The polynomial and Gaussian kernel based SVMs are designed to provide most optimized boundary for classification. The classification test results from SVMs are accurate for simulation model as well as experimental set-up, and thus provides fast and robust protection scheme for distance relaying in transmission line.

1 Introduction

This chapter presents machine learning technique such as Support Vector Machine (SVM) [1-10] for distance relaying of transmission line. SVM is a relatively new computational learning method based on the statistical learning theory. In SVM, original input space is mapped into a high-dimensional dot product space called a feature space, and in the feature space the optimal hyperplane is determined to

S.R. Samantaray
Department of Electrical Engineering,
National Institute of Technology Rourkela, Orissa, India

P.K. Dash
Centre for Research in Electrical, Electronics and Computer Engineering,
Bhubaneswar, Orissa, India

G. Panda
School of Electrical Sciences, IIT Bhubaneswar, Orissa, India
E-mail: sbh_samant@yahoo.co.in

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 171–198.
springerlink.com

maximize the generalization ability of the classifier. The optimal hyperplane is found by exploiting the optimization theory, and respecting insights provided by the statistical learning theory. SVMs have the potential to handle very large feature spaces, because training of SVM is carried out so that the dimension of classified vectors does not have as distinct influence on the performance of SVM as it has on the performance of conventional classifiers. That is why it is noticed to be especially efficient in large classification problems. This will also benefit in fault classification, because the number of features to be the basis of fault diagnosis may not have to be limited. Also, SVM-based classifiers are claimed to have good generalization properties compared to conventional classifiers, because in training the SVM classifier, the so-called structural misclassification risk is to be minimized, whereas traditional classifiers are usually trained so that the empirical risk is minimized.

A Support Vector Machine is a relatively new machine learning method that optimizes model on training data by solving a quadratic program (QP). In essence, an SVM finds the maximal separating hyperplane in feature space. It is computationally efficient because the transformation to feature space need not be done explicitly because dot-products in feature space can be represented by kernel functions.

SVM has advantages over traditional approaches such as neural networks for the following reasons:

1. Good generalization performance—once it is presented with a training set, it is able to learn a rule, which can correctly classify a new object quite often.
2. Computational efficiency—it is efficient in terms of speed and complexity.
3. Robust in high dimensions—in general, dealing with high-dimensional data is difficult for a learning algorithm because of over-fitting. One of the major reasons for attracting much attention is that SVMs are more robust to this over-fitting than other algorithms.

1.1 Support Vector Machine for Classification

Support Vector Machine (SVM) [1-10] is a relatively new computational learning method based on the statistical learning theory. In SVM, original input space is mapped into a high-dimensional dot product space called a feature space, and in the feature space the optimal hyperplane is determined to maximize the generalization ability of the classifier. The optimal hyperplane is found by exploiting the optimization theory, and respecting insights provided by the statistical learning theory.

SVMs have the potential to handle very large feature spaces, because training of SVM is carried out so that the dimension of classified vectors does not have as distinct an influence on the performance of SVM as it has on the performance of conventional classifiers. That is why it is noticed to be especially efficient in large classification problems. This will also benefit in fault classification, because sometimes the number of features to be the basis of fault diagnosis may be large. Also, SVM-based classifiers are claimed to have good generalization properties compared to conventional classifiers, because in training the SVM classifier, the so-called structural misclassification risk is to be minimized, whereas traditional classifiers are usually trained so that the empirical risk is minimized. SVM is

compared to the RBF neural network in an industrial fault classification task [11], and it has been found to give better generalization.

Let n-dimensional input $x\_i\,(i=1.....M), M$ is the number of samples) belong to class-I or Class II and associated labels be $y\_i=1$ for Class I and $y\_i=-1$ for Class II, respectively. For linearly separable data, we can determine a hyperplane $f(x)=0$ that separates the data

$$f(x)=w^{T}x+b=\sum_{j=1}^{n} w_j x_j+b=0 \tag{1}$$

where 'w' is an n-dimensional vector and 'b' is a scalar. The vector 'w' and the scalar 'b' determine the position of the separating hyperplane. Function *sign*(f(x)) is also called the decision function. A distinctly separating hyperplane satisfies the constraints $f(x_i)\geq 1$ if $y_i=+1$ and $f(x_i)\geq -1$, if $y_i=-1$. This result in

$$y_i f(x_i)=y_i(w^{T}x_i+b)\geq 1 \text{ for } i=1\ldots\ldots, M. \tag{2}$$

The separating hyperplane that creates the maximum distance between the plane and the nearest data, i.e., the maximum margin, is called the optimal separating hyperplane. An example of the optimal separating hyperplane of two datasets is presented in Fig. 1. From the geometry, the geometrical margin is found to be $\|w\|^{-2}$. Taking into account the noise with slack variables ξ_i and error penalty C, the optimal hyperplane can be found by solving the following convex quadratic optimization problem,

Minimize

$$\frac{1}{2}\|w\|^2+C\sum_{i=1}^{M}\xi_i$$

subject to

$$\begin{aligned} &y_i(w^{T}x_i+b)\geq 1-\xi_i, \text{ for } i=1.....M \\ &\xi_i\geq 0, \text{for all } i \end{aligned} \tag{3}$$

Where ξ_i is measuring the distance between the margin and the examples x_i lying on the wrong side of the margin. The calculations can be simplified by converting the problem with Kuhn–Tucker conditions into the equivalent Lagrange dual problem, which will be

maximize

$$W(\alpha)=\sum_{i=1}^{M}\alpha_i-\frac{1}{2}\sum_{i,k=0}^{M}\alpha_i\alpha_k y_i y_k x^{T}{}_i x_k$$

subject to

$$\sum_{i=1}^{M} y_i\alpha_i=0, C\geq\alpha_i\geq 0, i=1....,M \tag{4}$$

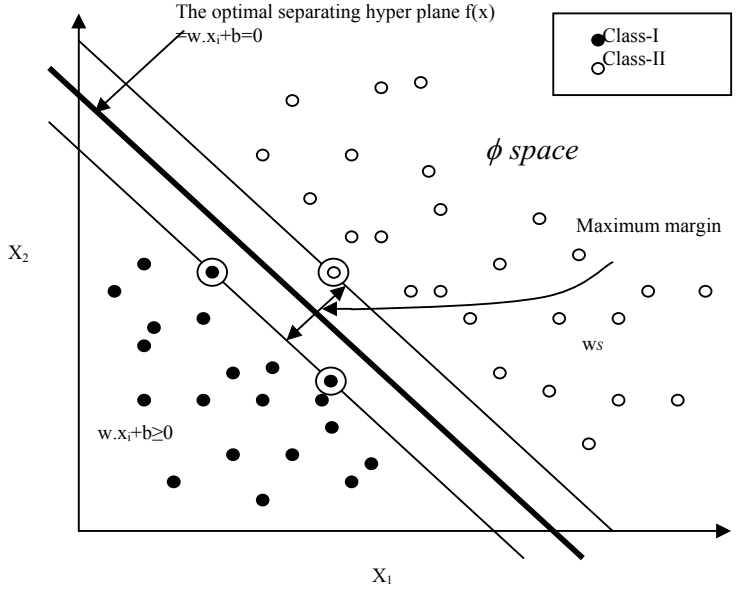


Fig. 1. f(x) as a separating hyperplane lying in a high-dimensional space. Support vectors are inside the circles.

The number of variables of the dual problem is the number of training data. Let us denote the optimal solution of the dual problem with α^* and w^*. According to the Karush–Kuhn–Tucker theorem, the inequality condition in (2) holds for the training input–output (feature and label) pair (x_i, y_i) only if the associated α^* is not 0. In this case the training example x_i is a support vector (SV). Usually, the number of SVs is considerably lower than the number of training samples making SVM computationally very efficient. The value of the optimal bias b^* is found from the geometry:

$$b^* = -\frac{1}{2}\sum_{SVs} y_i \alpha^*_i (s_1^T x_i + s_2^T x_i), \tag{5}$$

where s_1 and s_2 are arbitrary support vectors(SVs) for Class I and Class II, respectively. Only the samples associated with the SVs are summed, because the other elements of optimal Lagrange multiplier α^* are equal to zero.

The final decision function will be given by

$$f(x) = \sum_{SVs} \alpha_i y_i x_i^T x + b^* \tag{6}$$

Then unknown data example 'x' is classified as follows:

$$x \in \begin{cases} \text{Class} - \text{I} & \text{if } f(x) \geq 0 \\ \text{Class} - \text{II} & \text{Otherwise} \end{cases} \tag{7}$$

SVM can also be used in nonlinear classification tasks with application of kernel functions. The data to be classified is mapped onto a high-dimensional feature space, where the linear classification is possible. Using a nonlinear vector function $\phi(x) = (\phi_1(x).....\phi_m(x)), m >> n$ to map the *'n'*-dimensional input vector 'x' into the 'm' dimensional feature space, the linear decision function in dual form is given by

$$f(x) = \sum_{SVs} \alpha_i y_i \phi^T(x_i)\phi(x) \tag{8}$$

Working in the high-dimensional feature space enables the expression of complex functions, but it also generates problems. Computational problems occur due to the large vectors and the danger of overfitting also exists due to the high dimensionality. The latter problem is solved with application of the maximal margin classifier, and so-called kernels give solution to the first problem. Notice that in (8) well as in the optimization problem (3), the data occur only in inner products. A function that returns a dot product of the feature space mappings of original data points is called a kernel, $K(x,z) = \phi^T(x)\,\phi\,(z)$. Applying a kernel function, the learning in the feature space does not require explicit evaluation of ϕ. Using a kernel function, the decision function will be

$$f(x) = \sum_{SVs} \alpha_i^{*} y_i K(x_i x) \tag{9}$$

and the unknown data example is classified as before. The values of $K(x_i, x_j)$ over all training samples $i, j = 1.....M,$ from the kernel matrix, which is a central structure in the kernel theory. Mercer's theorem states that any symmetric positive-definite matrix can be regarded as a kernel matrix.

The polynomial learning machines of degree 'n' have the inner product kernel

$$K(x,z) = (x^T z + 1)^n \tag{10}$$

and radial basis function machines have the inner product kernel

$$K(x,z) = \exp\left\{-\frac{|x-z|^2}{2\sigma^2}\right\} \tag{11}$$

Where the 'σ' is the width of the Gaussian function.

2 Distance Relaying of an Advanced Series Compensated Transmission Line Using SVM

The use of FACTS [15-18] devices to improve the power transfer capability in high voltage transmission line is of greater interest in these days. The thyristor controlled series compensator (TCSC) is one of the main FACTS devices, which has the ability to improve the utilization of the existing transmission system. TCSC based compensation possess thyristor controlled variable capacitor protected by Metal Oxide Varistor (MOV) and an air gap. However, the implementation of this technology changes the apparent line impedance, which is controlled by the firing angle of thyristors, and is accentuated by other factors including the metal oxide varistor (MOV). The presence of the TCSC in fault loop not only affects the steady state components but also the transient components. The controllable reactance, the MOVs protecting the capacitors and the air-gaps operation make the protection decision more complex and, therefore, conventional relaying scheme based on fixed settings has its limitation. Fault classification and section identification is a very challenging task for a transmission line with TCSC. Different attempts have been made for fault classification using Wavelet Transform, Kalman filtering approach and neural network [19, 20].

The Kalman filtering approach finds its limitation, as fault resistance can not be modeled and further it requires a number of different filters to accomplish the task. Both BPNN (back propagation Neural Network), RBFNN (radial basis function neural network), FNN (Fuzzy Neural network) are employed for adaptive protection of such a line where the protection philosophy is viewed as a pattern classification problem [21, 22]. The networks generate the trip or block signals using a data window of voltages and currents at the relaying point. However, the above approaches are sensitive to system frequency-changes, and require large training sets and training time and a large number of neurons.

The research work presents a new approach for fault classification and section identification of TCSC based line using support vector machine (SVM). SVM, basically, is a classifier based on optimization technique. It optimizes the classification boundary between two classes very close to each other and thereby classifies the data sets even very close to each other. Also SVM works successfully for multiclass classification with SVM regression.

The current signals for all phases are retrieved at the relaying end at a sampling frequency of 1.0 kHz. Half cycle data (10 samples) and firing angle are used as input to the SVM. The SVM is trained with input and output sets to provide most optimized boundary for classification. Also another SVM is trained for identifying the TCSC position on the transmission line. Taking the current data samples before and after the TCSC, the corresponding SVM is trained to identify whether the fault includes TCSC or not. When fault includes TCSC, the 3^{rd} and 5^{th} harmonic components are highly pronounced compared to the fault which doesn't include TCSC. This issue is taken care by SVMs as the total half cycle (10 samples) data of the fault current signal is taken into consideration for training and testing the SVMs.

2.1 System Studied

A 440 kV, 50 Hz power system is illustrated in Fig. 2.In this system the TCSC is located at midpoint of the transmission line, used for the distance protection study. The power system consists of two sources, TCSC and associated components and a 300 km transmission line. The transmission line has zero sequence impedance Z(0)=96.45+j335.26 ohm and positive sequence impedance Z(1)=9.78+j110.23 ohm. $E_S = 400\,\text{kV}$ and $E_R = 400\angle\delta$ kV. The TCSC is designed to provide compensation varying form minimum 30% to maximum 40%. All the components are modeled using the EMTDC subroutines.

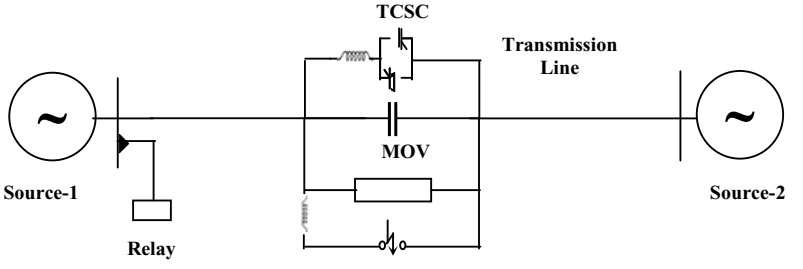


Fig. 2. The TCSC based line

The sampling frequency is 1.0 kHz at 50 Hz base frequency. The metal oxide varistor (MOV) consists of a number of zinc oxide disks electrically connected in series and parallel. The purpose of the MOV is to prevent the voltage across the capacitor from rising to levels which will damage the capacitor. This is most likely to happen when a fault occurs at a point on the compensated line which minimizes the impedance of the fault loop. When instantaneous voltage across the capacitor approaches a dangerous level the MOV begins to draw a significant proportion of the line current thereby limiting the voltage across the capacitor at that level. This action alters the impedance in the series path and hence the fault-loop impedance. In the event that the MOV remains in conduction long enough to raise its temperature (energy) to a dangerous level an air-gap is triggered to short out both the MOV and the capacitor, again changing the fault loop impedance. The operation of the MOV can be within the first half cycle of fault and depending on the severity of the fault, it may continue to operate until the air-gap is triggered cycles later. This is precisely the time when a digital relay makes protection decision. Further, a bypass switch in parallel with the gap automatically closes for abnormal system conditions that cause prolonged current flow through the gap. Fig. 3 shows the components of MOV and characteristics. The fault current variation with firing angle is shown in Fig. 4. The fault current pattern including TCSC and without including TCSC is shown in Fig. 5.

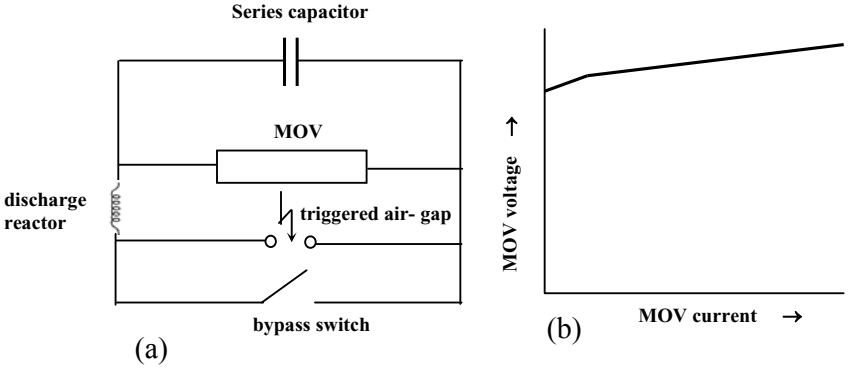


Fig. 3. (a) MOV protected series capacitor (b) MOV characteristic

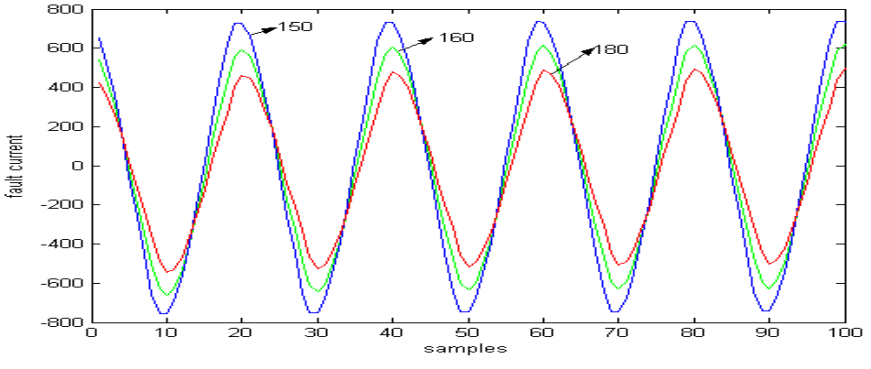


Fig. 4. Fault current with TCSC at different firing angles

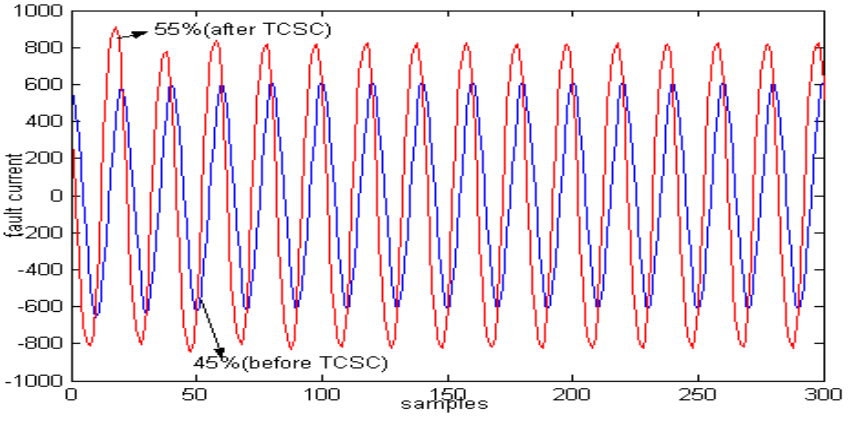


Fig. 5. Fault current before and after TCSC at 160º firing angle

The small inductance in the arrangement limits the current through the air-gap or switch circuit. The TCSC is designed such that it provides 30% compensation at 180^0 (minimum) and 40% compensation at 150^0 (maximum) firing angle and in this study the firing angle is varied within this range as shown in Fig. 6 .The proposed protection scheme is shown in Fig. 7.

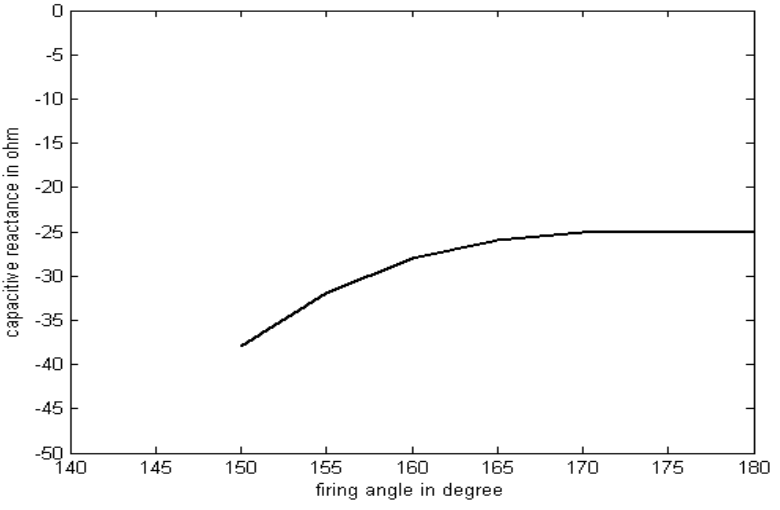


Fig. 6. Variation of capacitive reactance with firing angle

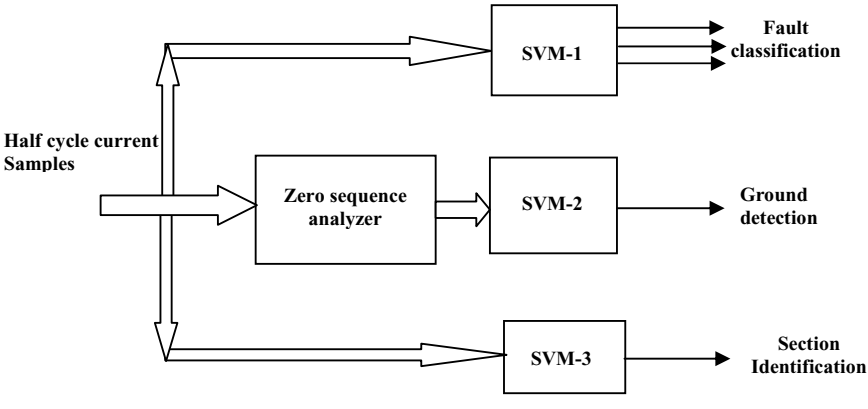


Fig. 7. Proposed scheme for protection. Fault classification (SVM-1), Ground detection (SVM-2) and section identification (SVM-3).

The TCSC is placed at 50% of the transmission line with 300 km line length, which is 150 km from relaying end. The simulation for all 11 types of shunt faults (L-G,LL-G,LL,LLL,LLL-G) are made on the transmission line with different fault resistance, source impedance, incident angles at different fault locations with varying the

firing angle from 150º-180º with (after) and without including(before) TCSC. The half cycle signal having 10 samples from the fault inception are retrieved at the relaying end and normalized to be used as input to the corresponding SVMs.

2.2 SVM Training and Testing

2.2.1 SVM for Fault Classification

The half cycle fault current signal samples after the fault inception are taken as input to the SVM. The corresponding output is either fault or no-fault condition. Ten samples(half cycle at 1.0 kHz sampling frequency) of fault current form the fault inception are retrieved at the relaying end are normalized along with the firing angle of TCSC and are used as input(11-inputs) space which is termed as 'x'. 'y' is the corresponding output which results '1' for fault and '-1' for no-fault condition. The optimal marginal classifier is designed with polynomial kernel with different order and Gaussian kernel with different parameter value. Both results are compared as depicted in Table-3.1.The SVM-1 is trained with 500 data sets and tested with 200 data sets, each set comprising of 11 data points(10 for half cycle current signal and 1 for firing angle of TCSC) for 'x' as input and (1,-1) for 'y' as corresponding output.

Table 1. Testing of SVM-1 for fault classification

Fault	Kernel	Parameter value	a	b	c
b-g fault at 30%,α=155º,R_f =20 ohm	Poly	n=2	-1	1	-1
	Poly	n=3	-1	1	-1
	Gaussian	σ=0.5	-1	1	-1
	Gaussian	σ =1.5	-1	1	-1
ab-g fault at 30%,α=165º,R_f =50 ohm	Poly	n=2	1	1	-1
	Poly	n=3	1	1	-1
	Gaussian	σ=0.5	1	1	-1
	Gaussian	σ =1.5	1	1	-1
'bc' fault at 45%,α=170º,R_f =100 ohm	Poly	n=2	-1	1	1
	Poly	n=3	-1	1	1
	Gaussian	σ=0.5	-1	1	1
	Gaussian	σ =1.5	-1	1	1
'abc' fault at 65%,α=160º,R_f =200 ohm	Poly	n=2	1	1	-1
	Poly	n=3	1	1	1
	Gaussian	σ=0.5	1	1	1
	Gaussian	σ =1.5	1	1	1
'abc-g' fault at 75%,α=165º,R_f =150 ohm with source changed	Poly	n=2	1	1	1
	Poly	n=3	1	1	1
	Gaussian	σ=0.5	-1	1	1
	Gaussian	σ =1.5	1	1	1

Table 2. Classification rates of SVM-1 for fault classification with 200 data sets

Fault	Kernel	Parameter Value	Classification rates (%)	No.of support vectors
L-G	Poly	n=2	96.52	15
	Poly	n=3	97.23	12
	Gaussian	σ=0.5	95.23	13
	Gaussian	σ =1.5	96.85	11
LL-G	Poly	n=2	96.27	9
	Poly	n=3	97.36	7
	Gaussian	σ=0.5	97.51	7
	Gaussian	σ =0.5	97.84	7
LL	Poly	n=2	96.84	11
	Poly	n=3	97.29	9
	Gaussian	σ=0.5	95.99	9
	Gaussian	σ =1.5	96.87	6
LLL	Poly	n=2	96.28	14
	Poly	n=3	97.56	12
	Gaussian	σ=0.5	95.68	11
	Gaussian	σ =1.5	96.87	10
LLL-G	Poly	n=2	97.25	12
	Poly	n=3	97.68	10
	Gaussian	σ=0.5	96.78	8
	Gaussian	σ =1.5	97.65	5

Faults on the line are simulated with various operating conditions including different incident angles, fault resistance (10-200ohm), source capacities, and various locations with different firing angles for all 11 types of shunt faults. When the parameter values of the polynomial kernel and Gaussian kernel are changed, the numbers of support vectors on the optimized marginal plane vary accordingly as seen from the result depicted in the Table-2. Here 'n' stands for the order of the polynomial and 'σ' stands for width of the gaussian function. The bound on the lagrangian multipliers 'C' is selected 10 and the conditioning parameter for QP method, lambda is chosen as 1.0*e-7. Different values of 'σ' with which the SVM is trained and tested are 0.5 and 1.5. Similarly the values selected for 'n' are 2 and 3. All the above parameters are selected after cross validation [12-14].

Table-1 shows the results for fault classification for various operating conditions. As seen from the table, for 'b-g' fault at 30%,α=155º,R_f =20 ohm, the 'b' ph output is '1' but output for 'a' and 'c' phases is '-1' for both polynomial and Gaussian kernel, which depicts that fault occurs only on 'b' phase. Also for 'abc' fault at 65%,α=160º,R_f =200 ohm, the output for all the phases is '1'. As seen, the misclassification occurs for the above operating condition with polynomial kernel with 'n'=2 resulting output of 'c' phase as '-1' instead of '1'.Table-3.2 depicts the classification rates at different faults and corresponding support vectors with polynomial and

Gaussian kernel of different parameter values. The classification rate is 95.23% (minimum) at L-G fault with Gaussian kernel with σ =0.5 and the support vectors are 13. Similarly the classification rate is 97.84% (maximum) fir LL-G fault with gaussian kernel with σ =0.5 which results 7 support vectors on the hyperplane.

2.2.2 SVM for Ground Detection

The ground detection is done separately by training another SVM. The peak value of the zero sequence component of the fault current signal for half cycle is found out for fundamental, 3rd and 5th harmonic component. The peak value of zero sequence components and firing angle of TCSC are used as the input-'x'(4-inputs) to the SVM-2 and the corresponding output(y) is '1' for the fault involving ground and '-1' for fault without involving ground. As the zero sequence components for these three harmonic components are pronounced in case of fault involving ground compared to fault without involving ground, the SVM-2 is trained to design a optimized classifier for ground detection.

Here 'n' stands for the order of the polynomial and 'σ' stands for width of the gaussian function. The bound on the lagrangian multipliers 'C' is selected 5 and the conditioning parameter for QP method, lambda is chosen as 1.0*e-7. Different values of 'σ' with which the SVM is trained and tested are 0.5 and 1.0. Similarly the values selected for 'n' are 1 and 2. All the above parameters are selected after cross validation [12-14].The SVM is trained with 500 data sets and tested for 200 data sets. The average classification rate for ground detection for 200 test cases is found to be 98.05% for all types of faults with different operating conditions. It is found form the Table-3 that for 'a-g' fault at 10%,α=160º,R_f =20 ohm, the output is'1' which shows that the fault involves ground. But 'bc' fault at 30%,α=165º,R_f =50 ohm, the output is '-1' which clearly shows that fault without involving ground. Also misclassification is observed for 'ac' fault at 45%, α=155º,R_f =200 ohm with polynomial kernel for n=2, which produces output '1' instead of '-1'. Also similar case happens for abc-g fault at 85%, α=160º,R_f =150 ohm with polynomial kernel for n=1.

2.2.3 SVM for Fault Section Identification

Fault section identification for the transmission line with TCSC is done by training the SVM-3 to build up an optimized classifier. The half cycle data (10 samples) after the fault inception and firing angle of TCSC are used as input-'x' (11-inputs) to the SVM and the output-'y' is the output. The output 'y' is '1' or '-1' for faults including TCSC and without TCSC, respectively. For any fault beyond 50% of the line the output of the SVM should be '1', otherwise '-1'.The SVM is trained with the bound on the lagrangian multipliers with 'C' selected as 20 and the conditioning parameter for QP method lambda chosen as 1.0*e-7. The lagrangian parameter 'C' is selected after testing the SVM with other values. The above parameters are selected after cross validation as mentioned earlier. The SVM is trained with 500 data sets and tested for 200 data sets. The average classification rate for section identification for 200 test cases is found to be 95.09% for all types of faults with different operating conditions.

Table 3. Testing of SVM-2 for ground detection

Fault	Kernel	Parameter value	Classification
a-g fault at 10%,α=160º,R_f =20 ohm	poly	n=1	1
	poly	n=2	1
	gaussian	σ =0.5	1
	gaussian	σ =1.0	1
bc fault at 30%,α=165º,R_f =50 ohm	poly	n=1	-1
	poly	n=2	-1
	gaussian	σ =0.5	-1
	gaussian	σ =1.0	-1
bc-g fault at 55%,α=175º,R_f =100 ohm	poly	n=1	1
	poly	n=2	1
	gaussian	σ =0.5	1
	gaussian	σ =1.0	1
abc fault at 65%,α=160º,R_f =150 ohm	poly	n=1	-1
	poly	n=2	-1
	gaussian	σ =0.5	-1
	gaussian	σ =1.0	-1
ac fault at 45%,α=155º,R_f =200 ohm	poly	n=1	-1
	poly	n=2	1
	gaussian	σ =0.5	-1
	gaussian	σ =1.0	-1
abc-g fault at 30%,α=165º,R_f =50 ohm	poly	n=1	1
	poly	n=2	1
	gaussian	σ =0.5	1
	gaussian	σ =1.0	1
bc-g fault at 85%,α=165º,R_f =100 ohm	poly	n=1	1
	poly	n=2	1
	gaussian	σ =0.5	1
	gaussian	σ =1.0	1
ab fault at 65%,α=160º,R_f =150 ohm	poly	n=1	-1
	poly	n=2	-1
	gaussian	σ =0.5	-1
	gaussian	σ =1.0	-1
abc-g fault at 85%,α=160º ,R_f =150 ohm with source changed	poly	n=1	-1
	poly	n=2	1
	gaussian	σ =0.5	1
	gaussian	σ =1.0	1

Table-4 depicts the results for section identification for TCSC on the transmission line. For 'ac-g' fault at 30%,α=165º,R_f =50 ohm, the output of SVM is '-1' which shows that the fault occurred before TCSC on the line. But for 'bc-g' fault at 55%,α=170º,R_f =100 ohm, the output of SVM is '1', which clearly depicts that

Table 4. Testing of SVM-3 for section identification

Fault	Kernel	Parameter value	Classification
ab fault at 10%, α=160º, R_f =20 ohm	poly	n=1	-1
	poly	n=2	-1
	gaussian	σ =0.5	-1
	gaussian	σ =1.0	-1
ac-g fault at 30%, α=165º, R_f =50 ohm	poly	n=1	-1
	poly	n=2	-1
	gaussian	σ =0.5	-1
	gaussian	σ =1.0	-1
bc-g fault at 55%, α=170º, R_f =100 ohm	poly	n=1	1
	poly	n=2	1
	gaussian	σ =0.5	1
	gaussian	σ =1.0	1
abc-g fault at 65%, α=170º, R_f =150 ohm	poly	n=1	1
	poly	n=2	1
	gaussian	σ =0.5	1
	gaussian	σ =1.0	1
ac fault at 45%, α=165º, R_f =100 ohm	poly	n=1	-1
	poly	n=2	-1
	gaussian	σ =0.5	-1
	gaussian	σ =1.0	-1
abc-g fault at 30%, α=175º, R_f =20 ohm	poly	n=1	1
	poly	n=2	-1
	gaussian	σ =0.5	-1
	gaussian	σ =1.0	-1
bc-g fault at 75%, α=165º, R_f =100 ohm with source changed	poly	n=1	1
	poly	n=2	1
	gaussian	σ =0.5	1
	gaussian	σ =1.0	1
ab fault at 15%, α=160º, R_f =20 ohm with source changed	poly	n=1	-1
	poly	n=2	-1
	gaussian	σ =0.5	1
	gaussian	σ =1.0	-1
abc fault at 65%,α=155º, R_f =200 ohm with source changed	poly	n=1	1
	poly	n=2	-1
	gaussian	σ =0.5	1
	gaussian	σ =1.0	1

the fault occurred after the TCSC on the line. Also misclassification is observed for 'abc-g' fault at 30%, α=175º,R_f =20 ohm with polynomial kernel with n=1 and for 'ab' fault at 15%,α=160º,R_f =20 ohm with source changed with gaussian kernel with σ =0.5. Also similar result occurs for 'abc' fault at 65%, α=155º,R_f =200 ohm with source changed for polynomial kernel with n=2.

3 SVM Based Distance Relaying for Single Circuit Transmission Line

In power transmission line protection, faulty phase identification and location of fault are the two most important items which need to be addressed in a reliable and accurate manner. Distance relaying techniques based on the measurement of the impedance at the fundamental frequency between the fault location and the relaying point have attracted wide spread attention. The sampled voltage and current data at the relying point are used to locate and classify the fault involving the line with or without fault resistance present in the fault path.

The accuracy of the fault classification and location also depends on the amplitude of the DC offset and harmonics in comparison to the fundamental component. Fourier Transforms, Differential equations, Waveform modeling and Kalman filters, and wavelet transforms are some of the techniques used for fault detection and location calculation [23-28]. Some of the recent papers in this area [25, 26, 28] have used only the sampled current values at the relaying point during faults for classification of fault types and distance calculations. To obtain more satisfactory results, however, wavelet filters having longer length and more levels of wavelet decomposition must be employed. Consequently more processing time is required: a fatal drawback for protection relays. The Kalman filtering approach finds its limitation, as fault resistance can not be modeled and further it requires a number of different filters to accomplish the task.

The speed and accuracy of distance relays of transmission lines can be improved by accurate and fast fault phase selection, and this also allows single pole tripping and autoreclosure to be employed. The selector module in the protective relay of a transmission line is very important as it is very much responsible for fault identification. So accurate fault detection is the primary requirement for protective relaying to start and trip correctly. In addition, faulty phase selection can be used to increase system stability by allowing single pole tripping and autoreclosure. Conventional approaches to phase selection are based on power frequency measurements suffer from deficiencies due to fault resistance, fault distance, influence of mutual coupling from adjacent lines, reactance effect, incomplete knowledge of system parameters etc. In this regard, some new techniques have been adopted. Approaches using travelling wave theory have been proposed to perform faulty phase selection. A method based on initial current travelling waves is presented in [29]. However, these approaches lead to increased hardware requirement. Travelling waves, being high frequency signals, are difficult to separate from interference noise.

In recent years, techniques using artificial neural networks (ANN) and fuzzy logic have been employed in faulty phase selection [30-32] due to their superior ability to learn and generalize from training patterns. However, in the fault classification and location tasks, the neural networks cannot produce accurate results due to the inaccuracies in the input phasor data and the requirement of a large number of neural networks for different categories of fault. BPNN (Back Propagation Neural Network), RBFNN (radial basis function neural network) and FNN (Fuzzy Neural network) are employed for adaptive protection of such a line where the protection philosophy is viewed as a pattern classification problem. The networks generate the trip or block signals using a data window of voltages and currents at the relaying point. However, the above approaches are sensitive to system frequency-changes, and require large training sets and training time and a large number of neurons.

This research work presents a new approach for faulty phase selection and ground detection using Support Vector Machine (SVM). A Support Vector Machine [1-10] is a relatively new machine learning method that optimizes model on training data by solving a quadratic program (QP). In essence, an SVM finds the maximal separating hyperplane in feature space. It is computationally efficient because the transformation to feature space need not be done explicitly because dot-products in feature space can be represented by kernel functions. Support vector machine (SVM)-based classification is a modern machine learning method that is surprisingly rarely used in fault classification even if it has given superior results in various classification and pattern recognition problems like, in text categorization [33] or phoneme recognition [34]. Currently there exist only a few publications concentrate on developing fault diagnostic methods based on SVM techniques.

The proposed method uses post fault current and voltage samples for $1/4^{th}$ cycle (5 samples) from the inception of the fault as inputs to the SVMs to result fault classification with ground detection. Polynomial and Gaussian kernel based SVMs are trained and tested with corresponding current and voltage samples to distinguish faulty phase from un-faulted one. SVM-1 and SVM-2 are designed to provide information regarding faulty phase and ground involved in the fault process respectively. The classification test results obtained form SVMs are accurate for simulation model with wide variations in operating conditions of the faulted power network, and thus provides fast and robust protection scheme for distance relaying.

3.1 System Studied

The power system model shown in Fig. 8 is simulated using PSCAD (EMTDC) software package. The relaying point is as shown in Fig. 8, where fault voltage and current signal samples are retrieved for different fault conditions. The power system network consists of two areas of 400 kV generation capacities and connected by 300 km long transmission line (distributed model).

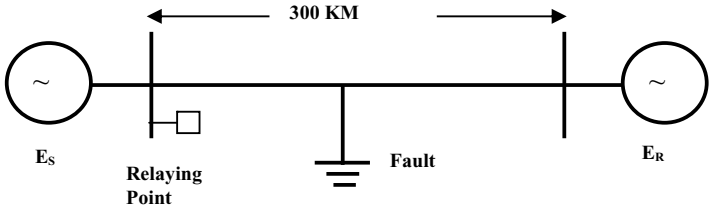


Fig. 8. Transmission Line Model

The transmission line parameters are as given below:

Zero sequence impedance (Z_{L0}) = 96.45 + j 335.26 ohms
Positive sequence impedance (Z_{L1}) =9.78 + j 110.23 ohms
Source impedances: $Z_S = 6 + j28.5$ ohms, $Z_R = 1.2 + j11.5$ ohms
Source voltages: $E_S = 400\,kV$, $E_R = 400\angle\delta\,kV$

where δ = load angle in degrees.

The power system model is simulated at 1.0 KHz sampling frequency. The voltage and current signals for different fault conditions are retrieved at the relaying point and fed to the SVMs for faulty phase selection and ground detection. The proposed SVM based relaying scheme is shown in Fig. 9.

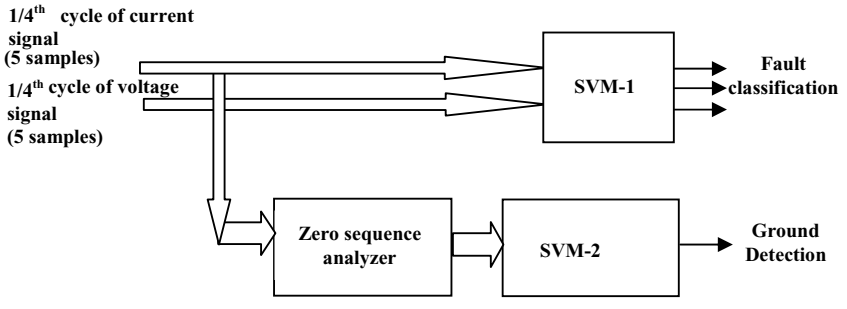


Fig. 9. Proposed scheme for protection. Fault classification (SVM-1), Ground detection (SVM-2)

The proposed algorithm is also tested on a physical transmission line model (experimental set up). The transmission line is consist of two 150 km π –section (total 300 km) and charged with 400 volt, 5 kVA synchronous machines at one end and 400 volt at the load end. The three phase voltage and current are stepped down at the relaying end with potential transformer (PT) of 400/10 V and current transformer (CT) of 15/5 A respectively. Data collected using PCL-208 Data Acquisition Card (DAC) which uses 12-bit successive approximation technique for A/D (Analog to Digital) conversion. The card is installed on a Personal Computer (P-4) with a driver software routine written in C++. It has 6 I/O channels with input voltage range of +/- 5 Volts. Data colleted with a sampling frequency of 1.0 KHz.

3.2 Computational Results

This section deals with the training and testing results obtained from the corresponding SVMs for faulty phase selection and ground detection. Both polynomial and gussian kernel based mapping are used to get the accurate results for classifying faults from un-faulted one. SVMs are trained with 300 data sets and validated for 300 data sets for each category of fault generated from simulation model and experimental set up separately. Initially, the Radial Basis Function Neural Network (RBFNN) is tested to have the performance comparison with the proposed SVM technique for fault classification.

3.2.1 Fault Classification Using Radial Basis Function Neural Network (RBFNN)

The fault classification results using Radial Basis Function Neural Network (RBFNN) [35] is given in Table-5. In this approach, a pruning strategy is used to select only a minimal number of hidden neurons by observing their outputs and if at any stage it is observed that the output of any neuron is insignificant it is omitted from the hidden layer. The current and voltage samples at the relaying point are retrieved for different fault condition and peak of the samples are fed to the RBFNN for fault classification. The RBFNN is tested with 300 datasets from simulation model. Table-1 provides the fault classification results using RBFNN for L-G, LL-G, LL, LLL-G, LLL faults. The classification rate is 95.12 % (maximum) in case of LLL-G fault and 93.75 % in case of LL fault.

Table 5. Classification rates of RBFNN for faulty phase selection

Fault	RBFNN Classification rates (%)
L-G	94.89%
LL-G	94.34%
LL	93.75%
LLL	94.84%
LLL-G	95.12%

3.2.2 Fault Classification Using Support Vector Machine

3.2.2.1 Phase Selection (SVM-1)

Initially 1/4th cycle fault signal is used for fault detection and the next 1/4th post fault current and voltage signal samples are used as inputs to the SVM. The corresponding output is either fault or no-fault condition. 5 samples of faulted voltage

and 5 samples of faulted current signal from the fault inception are retrieved at the relaying point and the corresponding normalized values are used as input features space (10 points) to the SVMs, termed as 'x_i'. The corresponding output is 'y_i' which results '1' for fault and '-1' for no-fault condition. The SVM-1 is trained with 300 data sets and tested with 300 data sets for each category of fault, each set comprising of 10 data points for 'x_i' as input and (1,-1) for 'y_i' as corresponding output.

Faults on the line are simulated with various operating conditions including different incident angles 'δ', fault resistance R_f (10-200ohm), source capacities and at various locations for all types of shunt faults. These shunt faults are Line-Ground (L-G) fault, Line-Line (L-L) fault, Line-Line-Ground (L-L-G) fault, Line-Line-Line (L-L-L) fault and Line-Line-Line-Ground (LLL-G) fault. In case of shunt faults, 'a-g', 'b-g', 'c-g' are categorized under L-G fault and 'ab-g', 'bc-g', 'ca-g' are categorized under LL-G fault. Similarly 'a-b', 'b-c', 'c-a' correspond to L-L fault and 'abc-g' corresponds to LLL-G fault. 'abc' is categorized under LLL fault. Thus there are 11 types of shunt faults occur on the power transmission line.

Here 'n' stands for the order of the polynomial and 'σ' stands for width of the gaussian function. The bound on the lagrangian multipliers 'C' is selected 5.0 and the conditioning parameter for QP (quadratic programming) method, lambda is chosen as 1.0*e-5. Different values of 'σ' with which the SVM is trained and tested are 1.0 and 1.5. Similarly the values selected for 'n' are 5 and 6. All the above parameters are selected after cross validation [12-14]. Different values of 'σ' and 'n' are used to make a comparison study on the classification rate and support vectors generated. When the parameter values of the polynomial kernel and Gaussian kernel are changed, the classification rate and the numbers of support vectors on the optimized marginal plane vary accordingly.

Table-6 provides the classification rates and support vectors during training of SVMs for faulty phase selection and ground detection. The classification rate for faulty phase selection is 99.26 %(maximum) with 14 nos of support vectors and for ground detection is 99.69 %(maximum) with 13 support vectors. After the SVMs are trained with the training data sets, the SVMs are tested or validated with test set data sets.

Table-7 shows the testing results for faulty phase patterns for various operating conditions. As seen from the table, for 'ab-g'(LL-G) fault at 30%, δ=45° °, R_f =50 ohm, the SVM-1 outputs for 'a' and 'b' are '1' but output for 'c' phase is '-1' for both polynomial and Gaussian kernel, which depicts that fault occurs on 'a' and 'b' phases. Also for 'bc' fault at 50%,δ=60° R_f =100 ohm the output for 'b' and 'c' phases are '1' but the output is'-1' for 'a' phase. As seen, the misclassification occurs for the 'abc-g' (LLL-G) fault at 90% ,δ=45° R_f =150 ohm with source changed, with gaussian kernel with 'σ'=1.0 resulting output of 'a' phase as '-1' instead of '1'.Table-8 depicts the classification rates at different

fault conditions with polynomial and Gaussian kernel of different parameter values during testing of SVM. The classification rate is 97.25% (minimum) at LLL-G fault with Polynomial kernel with n =5 and the classification rate is 98.87% (maximum) for LL-G fault with gaussian kernel with σ =1.0.

Table 6. Classification rates and support vectors during training of SVM-1 for phase selection and ground detection

Fault	Kernel	Parameter value	Training results for faulty phase selection		Training results for ground detection	
			Classification rates (%)	No. of support vectors	Classification rates (%)	No. of support vectors
L-G	poly	n=5	98.89	19	99.15	15
	poly	n=6	99.26	14	99.01	13
	gaussian	σ=1.0	98.11	15	99.00	12
	gaussian	σ =1.5	98.34	12	99.27	09
LL-G	poly	n=5	98.45	10	98.99	11
	poly	n=6	98.25	11	98.46	08
	gaussian	σ=1.0	98.99	09	99.49	10
	gaussian	σ =1.5	98.97	08	98.98	06
LL	poly	n=5	98.01	10	99.69	13
	poly	n=6	98.87	09	98.36	11
	gaussian	σ=1.0	98.20	08	98.77	09
	gaussian	σ =1.5	98.04	06	98.02	05
LLL	poly	n=5	98.32	19	99.03	17
	poly	n=6	97.98	14	98.00	14
	gaussian	σ=1.0	98.99	10	98.37	12
	gaussian	σ =1.5	98.03	09	98.53	09
LLL-G	poly	n=5	98.05	12	98.10	10
	poly	n=6	98.89	10	98.98	07
	gaussian	σ=1.0	97.99	08	99.18	09
	gaussian	σ =1.5	98.65	06	99.23	04

Table 7. Testing of SVM-1 for faulty phase patterns

Fault	Kernel	Parameter Value	a-phase	b-phase	c-phase
b-g fault at 10%,δ=30° R_f =10 ohm	poly	n=5	-1	1	-1
	poly	n=6	-1	1	-1
	gaussian	σ=1.0	-1	1	-1
	gaussian	σ =1.5	-1	1	-1
ab-g fault at 30%,δ=45° R_f =50 ohm	poly	n=5	1	1	-1
	poly	n=6	1	1	-1
	gaussian	σ=1.0	1	1	-1
	gaussian	σ =1.5	1	1	-1
'bc' fault at 50%,δ=60° R_f =100 ohm	poly	n=5	-1	1	1
	poly	n=6	-1	1	1
	gaussian	σ=1.0	-1	1	1
	gaussian	σ =1.5	-1	1	1
'abc' fault at 70%, δ=45° R_f =150 ohm	poly	n=5	1	1	-1
	poly	n=6	1	1	1
	gaussian	σ=1.0	1	1	1
	gaussian	σ =1.5	1	1	1
'abc-g' fault at 90% , δ=45° R_f =150 ohm with source changed	poly	n=5	1	1	1
	poly	n=6	1	1	1
	gaussian	σ=1.0	-1	1	1
	gaussian	σ =1.5	1	1	1
'ca-g' fault at 45%, δ=60° R_f =100 ohm	poly	n=5	1	-1	1
	poly	n=6	1	-1	1
	gaussian	σ=1.0	1	-1	1
	gaussian	σ =1.5	1	-1	1
'c-g' fault at 85%,δ=60°, R_f =150 ohm	poly	n=5	-1	-1	1
	poly	n=6	-1	-1	1
	gaussian	σ=1.0	-1	-1	1
	gaussian	σ =1.5	-1	-1	1
'ab' fault at 95%,δ=45°, R_f =200 ohm	poly	n=5	1	1	-1
	poly	n=6	1	1	-1
	gaussian	σ=1.0	1	1	-1
	gaussian	σ =1.5	1	1	-1
'abc-g' fault at 75%, δ=30°, R_f =200 ohm with source changed	poly	n=5	1	-1	1
	poly	n=6	1	1	1
	gaussian	σ=1.0	1	1	1
	gaussian	σ =1.5	1	1	1

Table 8. Classification rates of SVM-1 for faulty phase selection during testing

Fault	Kernel	Parameter Value	Classification rates (%)
L-G	Poly	n=5	98.16
	Poly	n=6	98.19
	gaussian	Σ=1.0	98.23
	gaussian	σ =1.5	97.89
LL-G	poly	n=5	98.15
	poly	n=6	98.09
	gaussian	Σ=1.0	98.87
	gaussian	σ =1.5	97.84
LL	poly	n=5	97.84
	poly	n=6	98.29
	gaussian	Σ=1.0	97.99
	gaussian	σ =1.5	97.87
LLL	poly	n=5	98.01
	poly	n=6	97.56
	gaussian	Σ=1.0	98.68
	gaussian	σ =1.5	97.87
LLL-G	poly	n=5	97.25
	poly	n=6	98.69
	gaussian	Σ=1.0	97.86
	gaussian	σ =1.5	98.15

3.2.2.2 Ground Detection (SVM-2)

The ground detection is done separately by training and testing SVM-2. The peak value of the zero sequence component of the fundamental ,3rd and 5th harmonic components of post fault current signal are found out and are used as the input-'x_i'(three input) to the SVM-2 and the corresponding output(y_i) is '1' for the fault involving ground and '-1' for fault without involving ground. As the zero sequence components are pronounced in case of fault involving ground compared to fault without involving ground, the SVM-2 is trained to provide the classification for ground detection. The SVM-2 is trained with 300 data sets and tested with 300 data sets for each category of fault, each set comprising of 3 data points for 'x_i' as input and (1,-1) for 'y_i' as corresponding output.

The order of the polynomial is n' and width of the gaussian function is 'σ'. The bound on the lagrangian multipliers 'C' is selected 5.0 and the conditioning parameter for QP method, lambda is chosen as 1.0*e-5. Different values of 'σ' with which the SVM is trained and tested are 1.0 and 1.5. Similarly the values selected for 'n' are 5 and 6. All the above parameters are selected after cross validation [12-14]. Table-9 provides the ground detection patterns for different fault conditions. It is found that for 'b-g' (L-G) fault at 10%,δ=30° R_f =10 ohm, the output is '1' which shows that the fault involves ground. But for 'bc' (L-L) fault at 50%, δ=60° R_f =100 ohm, the output is '-1' which clearly shows that fault does not involve ground. Also misclassification is observed for 'abc-g' (LLL-G) fault

Table 9. Testing of SVM-2 for ground detection patterns

Fault	Kernel	Parameter value	Ground detection pattern
'b-g' fault at 10%, δ=30° R_f =10 ohm	poly	n=5	1
	poly	n=6	1
	gaussian	σ=1.0	1
	gaussian	σ =1.5	1
'ab-g' fault at 30%, δ=45° R_f =50 ohm	poly	n=5	1
	poly	n=6	1
	gaussian	σ=1.0	1
	gaussian	σ =1.5	1
'bc' fault at 50%, δ=60° R_f =100 ohm	poly	n=5	-1
	poly	n=6	-1
	gaussian	σ=1.0	-1
	gaussian	σ =1.5	-1
'abc' fault at 70%, δ=45° R_f =150 ohm	poly	n=5	-1
	poly	n=6	-1
	gaussian	σ=1.0	-1
	gaussian	σ =1.5	-1
'abc-g' fault at 90% , δ=45° R_f =150 ohm with source changed	poly	n=5	1
	poly	n=6	1
	gaussian	σ=1.0	-1
	gaussian	σ =1.5	1
'ca-g' fault at 45%,δ=60° R_f =100 ohm	poly	n=5	1
	poly	n=6	1
	gaussian	σ=1.0	1
	gaussian	σ =1.5	1
'c-g' fault at 85%, δ=60°, R_f =150 ohm	poly	n=5	1
	poly	n=6	1
	gaussian	σ=1.0	1
	gaussian	σ =1.5	1
'ab' fault at 95%, δ=45°, R_f =200 ohm	poly	n=5	-1
	poly	n=6	-1
	gaussian	σ=1.0	-1
	gaussian	σ =1.5	-1
'abc-g' fault at 75%, δ=30°, R_f =200 ohm with source changed	poly	n=5	-1
	poly	n=6	1
	gaussian	σ=1.0	1
	gaussian	σ =1.5	1

at 75%, δ=30°, R_f =200 ohm with source changed, with polynomial kernel for n=5, which produces output '-1' instead of '1'. Also similar case happens for 'abc-g' (LLL-G) fault at 90%, δ=45° R_f =150 ohm with source changed, with

gaussian kernel for σ=1.0 which results '-1' instead of '1'. Table-10 shows the classification rate of the SVM-2 for ground detection. The classification rate is 99.32%(maximum) for LL-G fault with Gaussian kernel with σ=1.0 and the 97.89%(minimum) for LL fault with σ=1.5.The average classification rate for ground detection for 300 test cases is found to be 98.61% for all types of faults with different operating conditions.

Table 10. Classification rates of SVM-2 for ground detection during testing

Fault	Kernel	Parameter value	Classification rates (%)
L-G	poly	n=5	99.05
	poly	n=6	98.74
	gaussian	σ=1.0	98.98
	gaussian	σ =1.5	99.14
LL-G	poly	n=5	98.96
	poly	n=6	98.25
	gaussian	σ=1.0	99.32
	gaussian	σ =1.5	98.39
LL	poly	n=5	99.25
	poly	n=6	98.21
	gaussian	σ=1.0	98.69
	gaussian	σ =1.5	97.89
LLL	poly	n=5	98.47
	poly	n=6	97.98
	gaussian	σ=1.0	98.02
	gaussian	σ =1.5	98.07
LLL-G	poly	n=5	97.99
	poly	n=6	98.77
	gaussian	σ=1.0	99.05
	gaussian	σ =1.5	99.06

3.2.2.3 Phase Selection and Ground Detection for Experimental Data Sets

The proposed algorithm is also tested for data from experimental set-up. The trained SVMs are tested with 300 data sets from experimental set-up. The results for faulty phase selection and ground detection are given in Table-11. It is found that the trained SVMs produce accurate results when compared to the results using data from simulation model. The average faulty phase selection rate is 97.96% and the average ground detection is 98.50% compared to faulty phase selection and ground detection rates of 98.02% and 98.62% respectively using data from simulation study during testing.

Table 11. Classification rates of SVM-1 for faulty phase selection and SVM-2 for ground detection for experimental data sets

Fault	Kernel	Parameter value	Testing results for faulty phase selection	Testing results for ground detection
			Classification rates (%)	Classification rates (%)
L-G	poly	n=5	97.88	99.01
	poly	n=6	97.11	98.56
	gaussian	σ=1.0	96.98	98.76
	gaussian	σ =1.5	98.05	99.02
LL-G	poly	n=5	97.95	98.74
	poly	n=6	97.23	98.11
	gaussian	σ=1.0	98.23	99.25
	gaussian	σ =1.5	97.21	98.27
LL	poly	n=5	98.98	99.32
	poly	n=6	97.98	98.12
	gaussian	σ=1.0	98.43	98.56
	gaussian	σ =1.5	97.45	97.68
LLL	poly	n=5	98.65	98.32
	poly	n=6	97.12	97.78
	gaussian	σ=1.0	97.99	97.99
	gaussian	σ =1.5	97.76	98.01
LLL-G	poly	n=5	97.65	97.87
	poly	n=6	98.67	98.69
	gaussian	σ=1.0	99.01	98.97
	gaussian	σ =1.5	98.92	98.99

4 Conclusions

A new approach for the protection of Flexible AC Transmission Line with TCSC using machine learning technique such as support vector machine is presented in this research work. Half cycle post fault current samples and firing angles are used as input to the SVMs and the output is the corresponding classification. SVM-1 is used for fault classification, SVM-2 is used for ground detection and SVM-3 is used for section identification for the TCSC on the line, respectively. It is found that SVMs are trained to result most optimized classifier and with very less numbers of training samples compared to the neural network and neuro-fuzzy systems. Also the error found is less that 5% taking all SVMs to consideration. Hence the proposed method is very accurate and robust for the protection of transmission line including TCSC.

For transmission line without FACTs, 1/4th cycle post fault current and voltage samples are collected at the relaying point and fed to the SVMs as inputs and provides the information about the faulty phase and ground involved in the fault process. SVM-1 is trained and tested with the faulted voltage and current samples to provide faulty phase selection and SVM-2 is trained and tested with the peak of the zero sequence currents to result the ground involvement in the fault process respectively. The polynomial and Gaussian kernel based SVMs provide faulty phase selection and ground detection with error less than 2%. The results are compared with the RBF neural networks (previous work) and found better with respect to the efficiency and speed. The proposed method detects and classifies the faults within ½ cycle from the inception of fault (10 ms). Also the algorithm is tested for experimental set-up and provides accurate results for faulty phase selection and ground detection. Hence the proposed technique is fast, accurate and robust for the protection of large power transmission networks.

References

[1] Burges, C.J.C.: A tutorial on support vector machines for pattern recognition. Data Mining Knowledge Discovery 2(2), 121–167 (1998)

[2] Burges, C.J.C.: Geometry and invariance in kernel based methods. In: Schoelkopf, B., Burges, C.J.C., Smola, A.J. (eds.) Advances in Kernel Methods—Support Vector Learning, pp. 89–116. The MIT Press, Cambridge (1999)

[3] Chapelle, O., Vapnik, V.N.: Model selection for support vector machines. In: Solla, S., Leen, T.K., Muller, K.-R. (eds.) Advances in Neural Information Processing Systems, vol. 12, pp. 230–236. The MIT Press, Cambridge (2000)

[4] Bregler, C.: Learning and recognizing human dynamics in video sequences. In: Nevatia, R., Medioni, G. (eds.), Silver Spring, MD, pp. 568–574. IEEE Computer Society Press, Los Alamitos (1997)

[5] Chapelle, O., Vapnik, V.N., Bousquet, O., Mukherjee, S.: Choosing multiple parameters for support vector machines. Mach. Learn. 46(1-3), 131–159 (2002)

[6] Collobert, R., Bangio, S.: SVMTorch: a support vector machine for large-scale regression and classifcation problems. J. Mach. Learn. Res. 1, 143–160 (2001)

[7] Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, Cambridge (2000)

[8] Yajima, Y., Ohi, H., Mori, M.: Extracting feature subspace for kernel based support vector machines. Tech. rep., Technical Report 2001–5, Department of Industrial Engineering and Management, Tokyo Institute of Technology (2001)

[9] Mangasarian, O.L., Musicant, D.R.: Successive over relaxation for support vector machines. IEEE Trans. Neural Networks 10m, 1032–1037 (1999)

[10] Vapnik, V.N.: The nature of statistical learning theory. Springer, New York (1999)

[11] Ribeiro, B.: Support vector machines and RBF neural networks for fault detection and diagnosis. In: Proceedings of the 8th international conference on neural information processing, paper 191 (2001)

[12] Cawley, G.C., Talbot, N.L.C.: Fast exact leave-one-out cross-validation of sparse least-squares support vector machines. Neural Networks 17, 1467–1475 (2004)

[13] Joachims, T.: Estimating the generalization performance of a SVM efficiently. In: Proc. 17th international conf. Machine learning. Morgan Kaufman, San Francisco (2000)
[14] Kearns, M., Ron, D.: Algorithmic stability and sanity-check bounds for leave-one-out cross validation. In: Proc. 10th conf. comput. Learning theory, pp. 152–162. ACM, New York (1997)
[15] Helbing, S.G., Karady, G.G.: Investigations of an advanced Form of Series Compensation. IEEE Trans. on Power Delivery 9(2), 939–946 (1994)
[16] Larsen, E.V., Clark, K., Miske, S.A., Urbanek, J.: Characteristics and Rating Considerations of Thyristor Controlled Series Compensation. IEEE Transactions on Power Delivery 9(2), 992–1000 (1994)
[17] Noroozian, M., Angquist, L., Ghandhari, M., Anderson, G.: Improving Power System Dynamics by series connected FACTS devices. IEEE Trans. on Power Delivery 12(4), 1635–1641 (1997)
[18] Dash, P.K., Pradhan, A.K., Panda, G., Liew, A.C.: Adaptive relay setting for flexible AC Transmission Systems (FACTS). IEEE Trans. on Power delivery 15(1), 38–43 (2000)
[19] Song, Y.H., Johns, A.T., Xuan, Q.Y.: Artificial Neural Network Based Protection Scheme for Controllable series-compensated EHV transmission lines. IEE Proc. Gen. Trans. Dist. 143(6), 535–540 (1996)
[20] Song, Y.H., Xuan, Q.Y., Johns, A.T.: Protection Scheme for EHV Transmission Systems with Thyristor Controlled Series Compensation using Radial Basis Function Neural Networks. Electric Machines and Power Systems, 553–565 (1997)
[21] Song, Y.H., Xuan, Q.Y., Johns, A.T.: Protection of scheme for EHV transmission systems with thyristor controlled series compensation using radial basis function neural networks. Electric machines and power systems 25, 553–565 (1997)
[22] Song, Y.H., Johns, A.T., Xuan, Q.Y.: Artificial Neural Network Based Protection Scheme for Controllable series-compensated EHV transmission lines. IEE Proc. Gen. Trans. Disbn. 143(6), 535–540 (1996)
[23] Yaussef, O.A.S.: New Algorithms for phase selection based on wavelet transforms. IEEE transactions on Power Delivery 17, 908–914 (2002)
[24] Yaussef, O.A.S.: Online applications of wavelet transform to power system relaying. IEEE transactions on Power Delivery 18(4), 1158–1165 (2003)
[25] Yaussef, O.A.S.: Combined Fuzzy-Logic wavelet based fault classification for power system relaying. IEEE transaction on Power Delivery 19(2), 582–589 (2004)
[26] Osman, A.H., Malik, O.P.: Transmission line protection based on wavelet transform. IEEE transaction on Power Delivery 19(2), 515–523 (2004)
[27] Martin, F., Aguado, J.A.: Wavelet based ANN approach for Transmission line protection. IEEE transaction on Power Delivery 18(4), 1572–1574 (2003)
[28] Chanda, D., Kishore, N.K., Sinha, A.K.: A wavelet multi-resolution analysis for location of faults on transmission lines. Electric Power and Energy systems 25, 59–69 (2003)
[29] Boolen, M.H.J.: Traveling wave based protection of double circuit lines. Proceeding Inst. Elect. Engg. C 140(1), 37–47 (1993)
[30] Dalstein, T., Kluicke, B.: Neural Network approach to fault classification for high speed protection relaying. IEEE transaction, P&D 10(2), 1002–1011 (1995)
[31] Dash, P.K., Pradhan, A.K., Panda, G.: A novel Fuzzy Neural Network based distance relaying scheme. IEEE Transaction on Power Delivery 15(3), 902–907 (2000)

[32] Martin, F., Aguado, J.A.: Wavelet based ANN approach for Transmission line protection. IEEE transaction on Power Delivery 18(4), 1572–1574 (2003)
[33] Joachims, T.: Text categorization with support vector machines: Learning with many relevant features. Technical report, University of Dortmund (1997)
[34] Salomon, J.: Support vector machines for phoneme classification. M.Sc Thesis, University of Edinburgh (2001)
[35] Dash, P.K., Samantaray, S.R.: An accurate fault classification algorithm using a minimal radial basis function neural network. International journal of engineering intelligent systems 12(4), 205–210 (2004)

Power Quality

Kyeon Hur and Surya Santoso

Abstract. Providing power quality (PQ) for 21st-century needs is one of the widely accepted principal characteristics of the envisioned Smart Grid because we will have more and more PQ-sensitive loads such as microprocessor-based devices, critical manufacturing processes, and data centers [1]. Our future global competitiveness demands disturbance-free operation of the digital devices that empower the productivity of our economy[1]. It is expected that the Smart Grid will provide a reliable power supply with fewer and briefer outages, cleaner power, and self-healing power systems, through advanced PQ monitoring, analysis of the measurement data, diagnosis of PQ disturbances, identification of the root causes, and timely automated controls. Note that computational intelligence(CI) has been an integral, significant part of advancing and expanding the horizons of this PQ research. The capabilities and applications of CI for PQ are continually evolving due to advanced PQ monitoring (or recording) devices. The objective of this chapter is to present these CI applications such as signal processing and artificial intelligence (AI) techniques to help understand, measure, and mitigate PQ phenomena. This chapter also describes challenges and potential applications in turning raw PQ measurement data to a much more valuable and actionable knowledge in order to improve PQ performance and produce real financial benefits as well.

1 Introduction

Electric power quality (PQ) problems describe all electric power problems or disturbances in the supply system that prevent end-user equipment from operating properly. They encompass a wide range of different phenomena with time scales from tens of nanoseconds to steady state. In principle, there are four important groups of PQ disturbances frequently encountered in the distribution system [2]:

Kyeon Hur
Yonsei University, 262 Seongsanno, Seodaemun-gu, Seoul, 120-749, Korea
E-mail: khur@yonsei.ac.kr

Surya Santoso
The University of Texas at Austin, 1 University Station C0803, Austin, TX, USA
E-mail: ssantoso@ece.utexas.edu

[1] It is estimated that PQ disturbances of all types cost the U.S. economy at least $10 billion per year in lost productivity and damaged equipment [1].

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 199–234.
springerlink.com

- Momentary outages and voltage sag disturbances due to breaker operations to clear faults, and also due to poor voltage regulations
- Transient overvoltages due to energizing and switching of distribution system equipment, and lightning strikes entering the distribution system
- Harmonic distortion problems due to end-user nonlinear loads and the interaction of these loads with the distribution system equipment
- Small and gradual voltage (including voltage flicker leading to light flicker), current and frequency variations from the nominal values due to the inherent unbalanced nature of the distribution system.

PQ problems may originate in the distribution system (such as equipment energizing or circuit breaker clearing operations), and customer's own facility or other customer facilities (where their equipment produces electrical emissions – i.e., harmonic currents). They can also arise from the interaction among customer loads, or between customer loads and the utility system. It is also common that switching of distribution system equipment can interact with end-user equipment adversely. In addition to voltage sags and overvoltage phenomena, power system resonance phenomena arising from the interaction between distribution system equipment and end-users' loads[2] are a major important source of power quality disturbances. The interaction usually arises due to the shift in the system frequency response (caused by shunt capacitor bank energizing), the relatively weak system (i.e., low short-circuit capacity), and poor system damping. A real example of the power system harmonic resonance facilitated by capacitor bank energizing is illustrated in Fig. 1. Note that before the capacitor bank energizing, the harmonics are minimal, but it exhibits clear signs of harmonic distortion after the capacitor bank energization, shifting the system resonant frequency. Note also that transients are damped out rather slowly because it takes more than a quarter of a second, indicating low system damping. Impacts of power system resonance include additional losses, overheating, and overloading of capacitors, transformers, and motors. It can also cause interference with telecommunication lines and errors in power metering.

The majority of PQ problems can be characterized through measurements[3] of voltage and current [3, 4, 5]. Continuous monitoring over an extended period is often required to capture infrequent and unexpected disturbances. The monitoring is also used to track the ongoing system performance, to watch for conditions that could require attention, and to provide information for utility and customer personnel when there is a problem to be investigated. PQ monitoring has been widely used to evaluate system-wide performance as well. By understanding the normal power quality performance of a system, a utility can identify abnormal characteristics and offer information to customers to help them match their sensitive equipment characteristics with realistic PQ characteristics. For example, a technical understanding

[2] Widespread use of power electronic devices and the frequent energization of distribution system transformers may be considered to be main harmonic sources.

[3] Understanding how to monitor PQ conditions – i.e., understanding which types of measurements are needed, how to make them, over what time frame, and at what locations – is key to any PQ project.

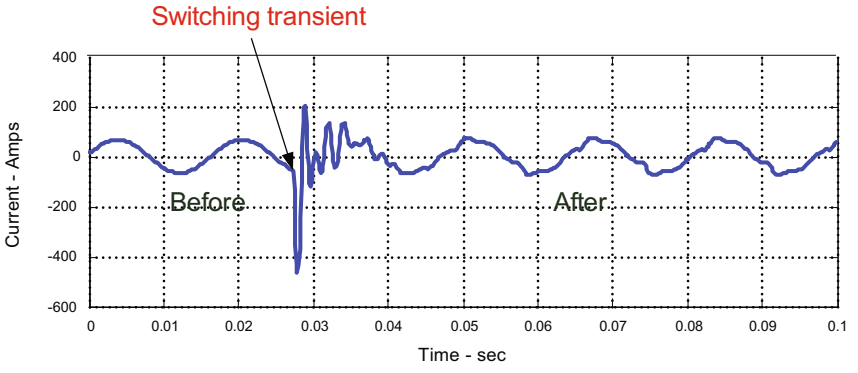


Fig. 1. Capacitor energization resulting in increased harmonic distortion

of waveforms and what shape the power needs to be in for a particular application is important to meet the operational requirements of end-use loads and various regulatory standards and guidelines.

Since the time scales of PQ disturbances vary widely, monitoring instruments such as PQ monitors, digital fault recorders, digital relays, and other intelligent electronic devices (IEDs) should ideally have the capability of capturing events ranging in frequencies from DC to a few megahertz for faster transients such as lightning and internal load switching[4]. Fig. 2 summarizes the basic requirements of monitoring as a function of the different types of PQ variations and subsequent analysis [4].

As presented in Fig. 2, the requirements for the PQ monitoring and analysis can be quite substantial. The PQ measurements are based on numerous metrics, including voltage regulation limits, voltage balance limits, harmonic distortion levels, voltage sag/swell rates, interruption rates, voltage flicker limits, switching and lightning transients, electrical noise, and unwanted stray voltage potentials [2]. It is very important that the user thoroughly understands the capabilities, limitations, and features of each type of PQ instrument in reference to the their needs. Unfortunately, as more and more PQ monitors are installed in the utility and customer facilities, end users of PQ monitors are often overwhelmed with voluminous data [4]. Note further that the aforementioned PQ phenomena may have a variety of different causes, which motivates more efficient tools or algorithms for improving the power quality and equipment performance. The manual acquisition and analysis of data has inherent inefficiencies and slows maintenance opportunities. It is desired to establish automated PQ data gathering, and sifting of the data that can streamline the PQ-related utility operations. A smart system to identify fundamental signatures of PQ events, correlate them to the root causes, and indicate impending equipment

[4] It is worth noting that many commercial power-quality monitoring instruments have sampling rates of 256 samples per cycle since the majority of PQ events – i.e., sags, swells, interruptions, flicker, and harmonics to the 50th – have frequency contents below 5 kHz because they are the majority of big PQ issues. However, note also that the high-end instruments with a higher sampling rate are required to analyze the faster transients [2, 4, 17].

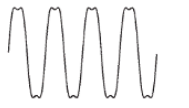
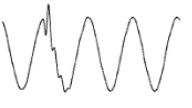
Type of power quality		Requirements for monitoring	Analysis and display requirements
Variation		▪ 3 phase voltages ▪ RMS magnitudes ▪ Continuous monitoring with periodic max./min./avg. samples ▪ Currents for response of equipment	▪ Trending ▪ Statistical evaluation of voltage levels and unbalance levels
Harmonic distortion		▪ 3 phase voltages and currents ▪ Waveform characteristics ▪ 128 samples per cycle minimum ▪ Synchronized sampling of all voltages and currents ▪ Configurable sampling characteristics	▪ Individual waveforms and FFTs ▪ Trends of harmonic levels (THD and individual harmonics) ▪ Statistical characteristics of harmonic levels ▪ Evaluation of neutral conductor loading issues ▪ Evaluation with respect to standards (e.g., IEEE 519, EN 50160) ▪ Evaluation of trends to indicate equipment problems
Voltage sags, swells, and short duration Interruptions		▪ 3 phase voltages and currents for each event that is captured ▪ Configurable thresholds for triggering events ▪ Characteristics of events with actual voltage and current waveforms, as well as RMS versus time plots ▪ RMS resolution of 1 cycle or better during the RMS versus time events and for triggering	▪ Waveform plots and RMS versus time plots with pre- and post-event information included ▪ Evaluation of cause of each event (fault upline or downline from the monitoring). ▪ Voltages and currents to evaluate load interaction issues ▪ Magnitude duration plots superimposed with equipment ride through characteristics (e.g., ITIC curve or SEMI curve) ▪ Statistical summary of performance (e.g., bar charts) for benchmarking ▪ Evaluation of power conditioning equipment performance during events
Transients		▪ 3 phase voltages and currents with complete waveforms ▪ Minimum of 128 samples per cycle for events from the power supply system (e.g., capacitor switching) ▪ Configurable thresholds for triggering ▪ Triggering based on waveform variations, not just peak voltage	▪ Waveform plots ▪ Evaluation of event causes (e.g., capacitor switching upline or downline from monitor) ▪ Correlation of events with switching operations ▪ Statistical summaries of transient performance for benchmarking

Fig. 2. Summary of monitoring requirements for different types of PQ disturbances

failure based on a robust knowledge base brings a great benefit [10, 18]. It should be emphasized that the true value of any power-quality monitoring program lies in its ability to analyze and interpret voluminous raw data and generate actionable information to prevent PQ problems or improve the overall power-quality performance. Signal-processing techniques in conjunction with various AI techniques are invaluable to meet this goal. The CI technology has enabled PQ monitoring and analysis to continuously become more powerful and to provide a growing array of benefits to the overall power-system operation and performance evaluation. The following sections provide an overview of how the CI techniques can be applied to provide better understanding and solve the various types of PQ problems, and they also discuss the techniques' unique functional features.

2 Computational Intelligence for PQ Applications

As indicated earlier, one of the important development areas for PQ is the implementation of CI techniques in PQ monitoring and analysis systems that can automatically

evaluate disturbances and operating conditions in order to draw conclusions about the cause of the problems or predict potential problems before they occur [5, 8]. It should be beneficial that the system can also contact the appropriate personnel and provide them with the relevant information using the advancing information and communication technologies [10]. Thus, numerous research activities have been focused on the following:

- Efficient algorithms that detect the instance of PQ disturbance and trigger recording of PQ data for further analysis
- Methodologies that not only detect PQ disturbance but also identify the PQ type by characterizing the disturbance to enable timely and appropriate control actions
- Tools for correlating the PQ disturbance with the root cause
- Methodology for locating the source of the PQ disturbance and system equipment that facilitates the PQ disturbance
- Techniques for efficiently managing the huge amount of data collected through the PQ monitors by storing only key signatures of important PQ disturbances.

To effectively discuss and better understand the integral roles of signal processing and AI-based techniques in PQ applications, let's first look into an envisioned intelligent PQ monitoring and analysis system[5] designed especially to deal with a variety of capacitor switching transient concerns [7, 10], as depicted in Fig. 3. It includes the PQ event detection and classification module with five functional modules, which autonomously examine the location the capacitor bank, proper capacitor switching, switching transient magnification, overall capacitor bank health, and system harmonic resonance.

This envisioned smart system covers various functional capabilities that existing research in the PQ area strives for, as stated above. Specifically, capacitor bank maintenance has been a time-consuming task for utility engineers. Typically, capacitor bank problems are found during routine inspections of capacitor banks or when failures result in outages or other system problems. The delay in finding problems such as blown fuses or can failures can result in inefficient system operation, voltage control issues, voltage unbalance, and harmonic problems. Problems such as switch restriking, harmonic resonance, and magnified transient voltages can cause catastrophic failures if they are not detected and fixed. An automated system can detect many of these problems before they cause catastrophic failures and can reduce time to repair routine problems that cause voltage-control and power-quality problems [10]. It is important to note that this research motivation and the values can be easily extended to other important PQ disturbances such as voltage-sag problem [6].

As discussed earlier, many utilities have already implemented extensive power-quality monitoring systems with continuous management of the monitoring databases [4]. These systems provide the data that can be used for automated identification of

[5] This may be a good illustrative example of advanced smart PQ monitoring and analysis systems. Implementation and development of the individual functional modules have been the active research topics in the PQ area for many years, and remarkable progress has been made. The performance and efficacy are evolving, and efforts to resolve the interoperability and integration issues are noticeable.

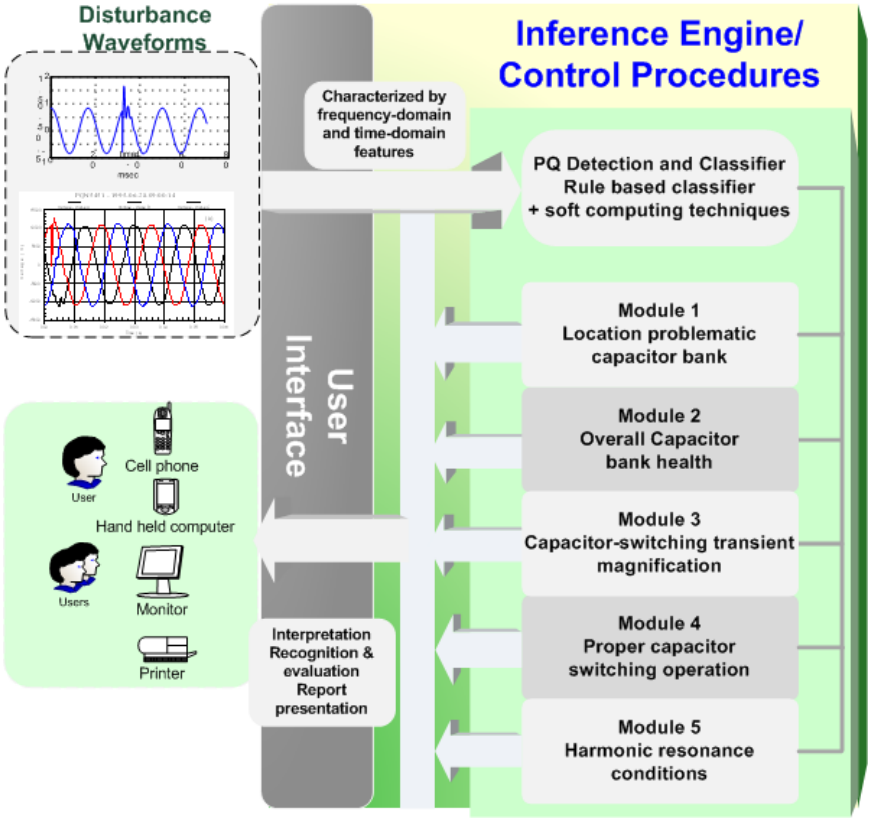


Fig. 3. A scalable expert system for capacitor switching transient concerns

PQ events. The data include transient voltages and currents during switching operations, as well as steady-state trends of voltage, current, reactive power, and harmonics. They are the valuable (but huge) data sets used to detect and analyze PQ concerns. Advanced signal-processing techniques based on the expertise of PQ professionals help detect the event and capture unique features of a certain PQ event automatically. These processes can be used in both time and frequency domains, and the techniques include classic but efficient FFT-based methods [20], wavelet transforms [8, 9], S-transform [11, 12], time-frequency analysis [13], and parametric methods such as Prony analysis, MUSIC, ESPRIT [6], etc. It is worth noting that different PQ disturbances require different signal-processing techniques, and sophisticated methods do not always perform best. Computationally simple, steady-state trend analysis of time-series data performs excellent for detecting and analyzing the voltage sag. Using the identified features, one can further categorize the PQ disturbances into the predetermined PQ types (and may even create a new type) with the help of AI-based tools such as expert system, soft computing techniques like neural network, fuzzy logic, evolutionary computation, support vector machine (SVM) or a combination

thereof [16, 6]. Heuristics based on engineers' practical knowledge and the underlying physics of the PQ phenomena reinforce the performance and may also provide computationally simple solutions. In fact, computational efficiency is especially required so that feasible and low-cost solutions for PQ monitoring equipment implementation may be achieved. Identifying impending failure of system equipment (capacitor bank, transformer, recloser, etc.) through equipment health monitoring should be greatly beneficial, as shown in the illustrative example. It will also be beneficial if the intelligent system can (1) correlate the PQ disturbance with the system equipment facilitating the disturbance by determining its location, and (2) evaluate power system operating conditions (abnormal switching condition and system resonance)and system characteristics (damping, resonant frequencies) continuously [10].

The advanced PQ monitoring and analysis system enables a preventive and proactive maintenance strategy as well. First, when forewarned of an imminent equipment failure, the utility has time to go and repair or replace the system before outages result. Second, it saves money by warning the utility of ailing equipment before it actually dies. Often this equipment can be repaired or relocated to keep it in operation. Third, this system gives the utility greater knowledge about various PQ events occurring on the grid. The upshot of all this is improved power quality. Such an automated, intelligent PQ evaluation system could quite possibly save utilities considerable money and shed new light on old PQ problems.

In the following sections, the various aforementioned CI techniques applied to improve the PQ applications are discussed. Applications include PQ detection and characterization, PQ classification, locating the correlated system equipment, power system characterization, and data compression.

2.1 PQ Analysis – Detection and Characterization of PQ Problems

To identify whether PQ is acceptable or not, we must have an understanding of the range of disturbances that can be present on the power system, the standards by which these disturbances are judged, and the expected impacts on loads and system equipment. Therefore, a good understanding of the various standards and guidelines is a prerequisite for the PQ analysis. For example, there are Institute of Electrical and Electronics Engineers (IEEE) and American National Standards Institute (ANSI) standards that specify the allowable range of conditions for total harmonic distortion and the specific levels of harmonics that can be present (IEEE Std. 519-1992). There exist specifications for voltage regulation and voltage balance (ANSI Std. C84.1-1995). IEEE Std. 1159-1995 discusses all of the various PQ disturbance types and provides a comprehensive manual about how to measure and guide them. For voltage sags and interruptions, the industry standards do not strictly define the specific rates or severity of these disturbances because the performance needs are more subjective and are interpreted based on load criticality, the susceptibility of the load to such disturbances, and power system design factors. Obtaining a working knowledge of the expected interruption and sag rates for various types of power systems in typical operating environments should be beneficial for planning and evaluating PQ studies [17].

2.1.1 Detection of PQ Problems

The analysis of a PQ problem begins with its detection [3, 8, 14, 9, 15]. The detection also provides the instant or triggering point in reference to the time when the PQ disturbance starts. The threshold values for triggering are set judiciously for different types of PQ disturbances of interest[6]. The values may be root-mean-square(rms) magnitudes of voltage and current, peak values of the waveforms, total harmonic distortion(THD) like harmonic indices, percent unbalance of voltages, or magnitudes of individual harmonic components, to name a few. These threshold values could be set and represented most effectively in the time, frequency, time-scale, or time-frequency domains, depending on certain PQ characteristics and applied techniques. It is also worth noting that the PQ event detection often triggers the data recording for further analysis. Although there may be some latency issues, modern GPS-based PQ monitors can provide accurate time stamps for the monitored data, which allows users to analyze multiple sets of PQ data associated with the same event from different monitoring sources simultaneously and to perform more advanced studies such as finding the event location, an impending equipment failure or the root cause of an event, etc. [4].

The end time of the event should be beneficial for further analysis. It is interesting to note that the start and end timings for a voltage dip may easily be detected because it generally results in a sudden drop and rise, as well as slow decay of the voltage magnitude between the instants. This is based on the characteristics of the voltage dip [3, 9]. However, the end time of a certain PQ event such as an oscillatory PQ event may not be as clear as the start point. In this case, heuristics based on extensive monitoring and analysis experience may help determine the transient- and steady-state (quasi-stationary) periods for further analysis. These transient and steady-state periods are often called transition segments (during an event) and event segments (both before and after an event). The start and end points are closely related to these segments [6]. For example, let's revisit a capacitor bank energizing transient case, as shown in Fig. 1. Prior to the energization, the total harmonic distortion is minimal, and it does not exhibit any sort of resonance problems. However, when a capacitor bank is energized, one can observe the transient period where the current is damped out, taking more than a quarter of a second because the system is relatively weak. After the transient period, the current waveform exhibits clear signs of harmonic distortion due to the resonance. Thus, once the capacitor bank energizing transient event occurs and the start point is accurately identified, one may extract the transient (data for less than a cycle after the event) and steady-state segments (before and after the event, allowing a few cycles of data) in an empirical but an automated manner [20].

[6] The objective is to be just sensitive enough to capture all of the significant PQ disturbances that are pertinent to the study, but not so sensitive that the instrument constantly triggers on the thousands of insignificant events down near the noise level that occurs daily. This also keeps the data build-up manageable by avoiding premature filling up of the memory [17].

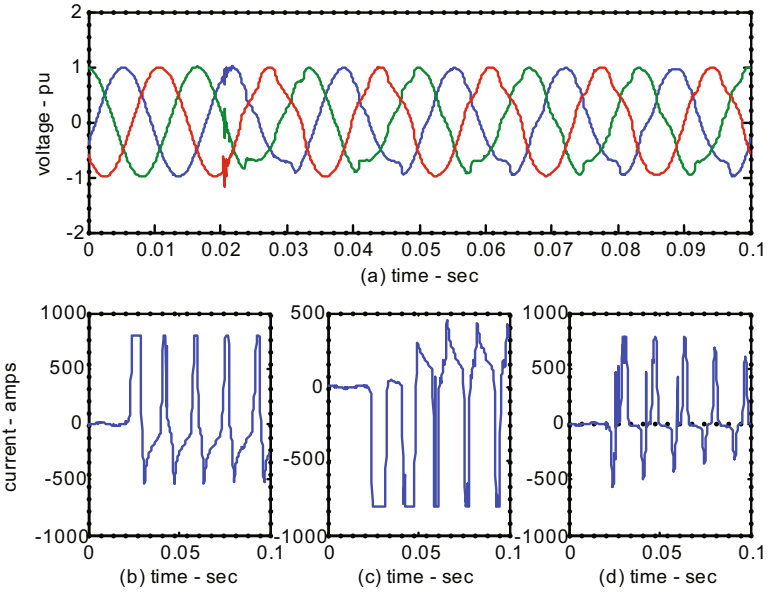


Fig. 4. A dynamic overvoltage condition due to transformer energizing

Dynamic overvoltage due to transformer energizing, as illustrated in Fig. 4, is another good example, and it is worth noting because the segment after the transient period may not be considered quasi-stationary, unlike the general power system harmonic problems. As well known, energizing the transformer produces magnetizing inrush current that can excite the power system's natural response frequencies [21]. It is apparent in Fig. 4 that the overvoltage magnitude is approximately 1.5 p.u. and that it persists over the entire monitoring record. Note that the inrush current peaks for the first few cycles are very large and highly distorted. Note also that the magnitude of the inrush current even exceeds the monitoring current setting of 200 amps, and thus the current waveshape appears flat-topped. Due to the randomly fluctuating transient flux in the iron core, it is typical to observe that one phase experiences a more severe inrush than others. This also causes current and thus voltage imbalances. Hence, it may require at least several cycles of PQ data for further analysis to classify this event among the similar types of PQ events.

Methods of PQ Detection

The underlying idea of PQ detection lies in finding a deviation from the quasi-stationary characteristics of voltage and current measurements from the monitoring devices [3, 6]. It is to find the transition segments, and three well-known approaches exist for detecting these transition segments, as follows:

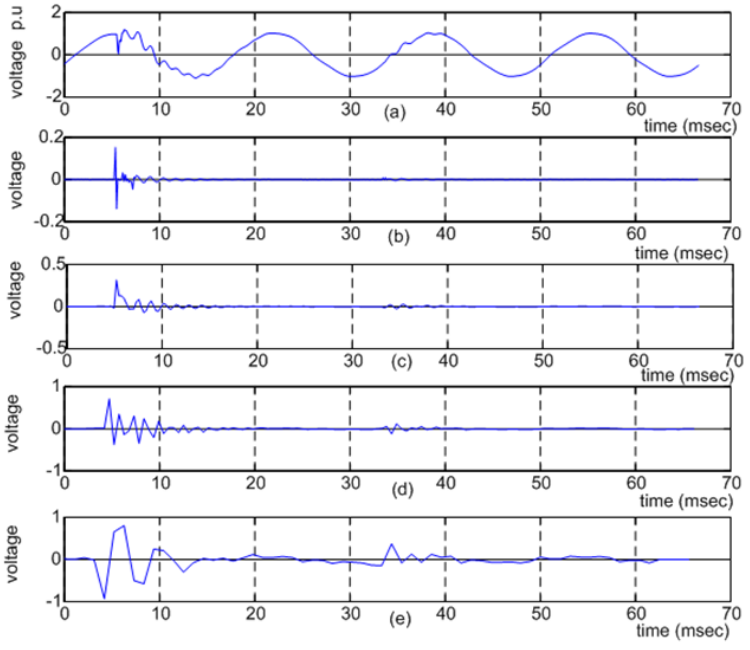


Fig. 5. Wavelet transform coefficients of a capacitor switching transient

- Calculation of time-dependent rms voltage and current magnitudes. This is the simplest but also a very efficient method for detecting many of the PQ disturbances such as voltage sags, swells, and interruptions. The transition segments are detected by comparing the change in magnitude of the calculated rms values with a predetermined threshold [2].
- Application of high-pass, band-pass filters and wavelet filters. This method is especially useful for detecting PQ disturbances that result in fast changes or oscillations. Numerous studies using wavelet transforms are noticeable in the literature because wavelet transforms are effective in detecting multi-scale singular points (sudden changes), and because transient segments are closely related to these singularities of the signal waveforms. Generally, some postprocessing is required to identify the start and end points of the transition segment because wavelet-transformed signals show all multi-scale singular points of a signal waveform. Fig. 5 illustrates a voltage waveform of a capacitor bank energizing and the squared wavelet transform coefficients at scales 1 – 4. In performing wavelet transformations, a given signal is decomposed into four decomposition scales with a Daub4 wavelet as a mother wavelet. A Daub4 wavelet is a Daubechies wavelet with 2 vanishing moments [3].
- Use of parametric signal models such as a damped sinusoidal model or autoregressive model. Depending on the algorithms used, a recorded data sequence

may be divided into blocks, and the model parameters in each block may be estimated. This can be accomplished using estimation of signal parameters via rotational invariance techniques (ESPRIT), multiple signal classification (MUSIC), or autoregressive (AR) modeling. Alternatively, iterative algorithms may be used without dividing data into blocks (e.g., using Kalman filters [33]). The so-called residuals (model errors), which indicate the deviation between the original waveform and the waveform generated by the estimated model, are then calculated. As long as the signal is quasi-stationary, the residual is small; however, for a sudden change in the signal, – e.g., a transition – the residual values become large. Residual values can therefore be used to detect transition segments [6].

Each of these methods has advantages and disadvantages. It is desired to evaluate and compare the performance of different methods for actual PQ disturbance data in terms of their ability to detect and localize the changes of the measured waveforms. The metrics may include time resolution, the detection rate, and the false alarm rate of the detected points.

2.1.2 Characterization of PQ Problems

PQ phenomena have been investigated directly from actual recorded disturbance waveforms thanks to widely available power monitor equipment. These disturbance recordings are stored as three-phase voltage and current time-series. They contain a wealth of information about the characteristics of the associated PQ types. Since these characteristics are unique to each PQ event, they are useful to portray the process of PQ disturbances and to pinpoint their root causes. The Fourier transform has been used to characterize steady state phenomena, whereas the wavelet transform has been applied to transient phenomena. We also make use of other mathematical analysis techniques such as curve-fitting techniques to estimate sudden voltage step changes.

PQ events are generally characterized by their maximum amplitudes, crest voltages, rms, frequency, statistics of wavelet transform coefficients, instantaneous voltage drops, number of notches, duration of transients, etc. These characteristics are different for each power quality event, thus they are unique identifying features. In the following, for example, we will classify capacitor switching transient events as normal energizing and back-to-back energizing events. We will further discuss characteristics of each category and formulate a procedure to extract identifying features from input data [3].

Energizing a capacitor bank is accompanied by an inrush current that flows into the bank. The characteristic impedance of the power system and the resistance of the line initially limit the inrush current. During energization, a sharp reduction in the voltage is immediately followed by a voltage rise, which later decays at a frequency determined by the inductance and capacitance of the circuit. Typically, the voltage rise due to capacitor switching operation can rise from 1.1 to 1.4 times the normal voltage. The oscillation frequency in the voltage waveform is between 300-1000Hz and lasts for less than half a cycle of the power frequency [3].

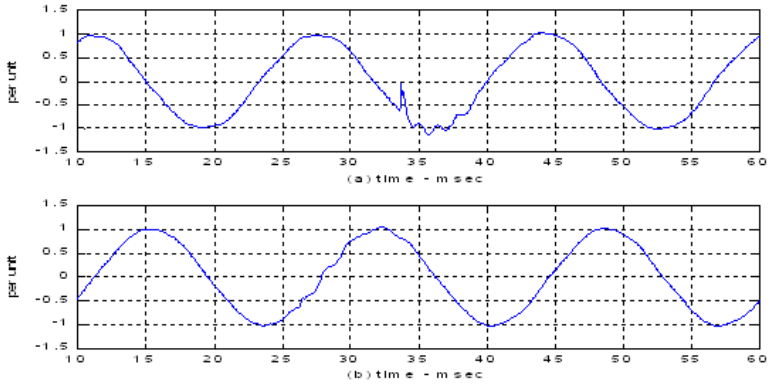


Fig. 6. Typical voltage waveforms for normal capacitor switching events. PQ monitor is located relatively (a) near and (b) faraway from the capacitor bank.

Normal Energizing of Utility Capacitors

Energization of utility capacitors is a daily operation in the utility system. They are switched into the system in anticipation of load increase at a customer site, to correct power factor, to support voltage on the system, and for many other reasons as well. These are all considered normal energizations and they each have their own unique features. The following are signatures for normal capacitor switching:

- Overvoltage: At the switching instant, the voltage in the capacitor cannot change instantaneously. The bus voltage is pulled down, and then rises as the capacitor begins to charge. During this process, the capacitor voltage may overshoot and ring at the natural frequency. The overvoltage under normal energizing is usually from 1.1 to 1.4 p.u. Fig. 6 shows the voltage waveform of normal energization of a utility capacitor.
- Polarity and Magnitude of the Step Voltage: One of the most common identifying features for normal energizing of utility capacitors is the polarity of the step voltage. If the power quality monitor is located at or near capacitors that have no series reactor (i.e., inrush reactor), a fast initial voltage step will be observed. The voltage step at the instant of closing cannot go beyond zero if the capacitor is lacking an initial net charge at the closing instant or if the capacitor is grounded. If the power quality monitor is located farther away from the capacitor, the voltage step change may not be observed or may not be as prominent. Both of these situations are illustrated in Fig. 6(a) and Fig. 6(b), respectively. In any event, sudden changes of voltages never go across the zero line (i.e., they do not change polarity). This behavior is exhibited in nearly all utility capacitors energizing.
- Oscillation Frequency of the Energizing Event in the Phase Voltage: The oscillation frequency of any kind of capacitor energizing in the phase voltage is generally between 300-1000 Hz. Thus, the frequency of oscillation is helpful in identifying

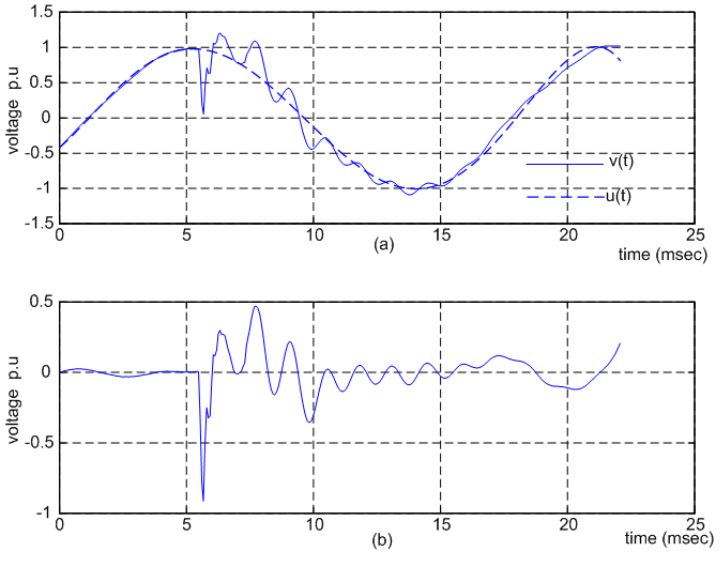


Fig. 7. Oscillation frequency of capacitor energizing

capacitor energizing in general, but it cannot be used to discriminate normal energizing from other types of capacitor energizing. Extracting the oscillation frequency from the capacitor energizing transients is difficult. First, the transient event from a given voltage waveform has to be localized using the squared wavelet transform coefficients. Once it is localized, the sinusoidal part has to be removed. Then, the frequency of the transient event in the voltage phase is now estimated by counting the number of crest voltages that are larger than 0.1 p.u. and the duration of the corresponding crest voltages. The solid line waveform shown in Fig. 7 (a) is the first 22 ms of phase C voltage in a typical normal capacitor switching event and the dotted line waveform is a curve-fitted waveform $u(t)$. The transient event can now be removed from $v(t)$ by subtracting $u(t)$ from it. The removed waveform $r(t)$ is now shown in Fig. 7 (b). The method does not always work if the removed waveform is very small in magnitude. In such a case, the oscillation frequency characteristic from the phase voltage cannot be extracted.

- Wavelet Transform Coefficients (WTC): The wavelet transform is utilized to provide additional features in identifying capacitor switching events. It is used to decompose the signal in different frequency bands and study its characteristics separately. Many different types of wavelet transforms have been applied to identify power system transient events. In developing a recognition and identification module, many prominent wavelet-based transforms will be tested and enhanced to allow identification of additional features of capacitor switching transient events. In fact, the aforementioned Fig. 5 shows wavelet transform coefficients for a typical capacitor switching transient event of phase C with scales from 1 to 4, (b) through (e), respectively. In identifying normal energization of

utility capacitors, the maximum squared magnitude with scales of 1 and 2 can indicate the onset of the disturbance and its distribution activity, respectively. The duration of the transient can also be estimated using WTCs at scales 1 and 2. The small oscillation wave following the spike-like event in WTCs at scales 1 and 2 provides an estimate of the engineering transient duration.

- Oscillation Frequency of the Energizing Event in the Phase Current: As described before, the oscillation frequency of the phase voltage transient is helpful in discriminating capacitor energizing events from other events. However, the frequency signature from the phase voltage cannot be used to distinguish between normal and irregular capacitor energizations. Fortunately, the oscillation frequency of the transient event at the phase current of a normal energizing is very distinct from other types of capacitor energizing disturbances. The frequency of the phase current during energization is typically between 300 and 1000 Hz. Once the transient event is localized by the wavelet transform, the oscillation frequency is computed using a Fourier transform. Current waveforms for a typical capacitor switching transient event are shown in Fig. 8.

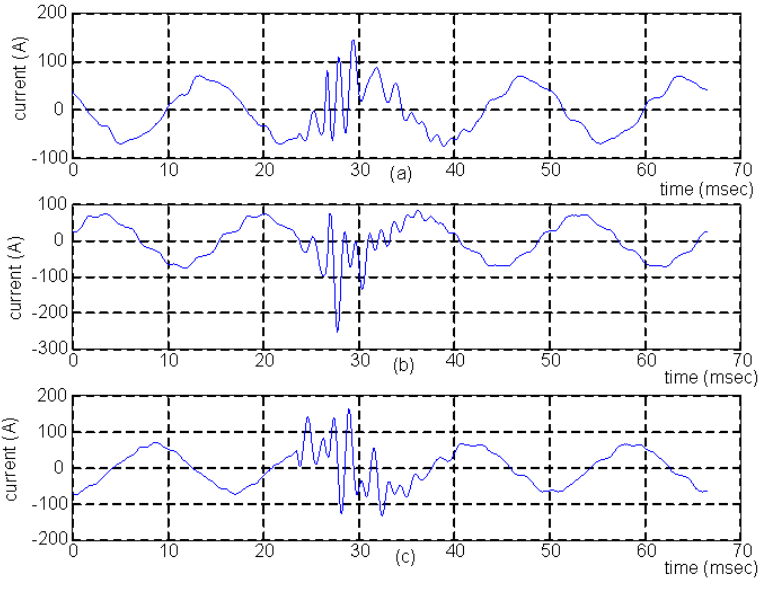


Fig. 8. Current waveforms at phases A, B, and C for a typical capacitor switching transient

Back-to-Back Capacitor Energizing

Back-to-back energizing of utility capacitors occurs when the capacitor being switched into the utility system is located near an already energized capacitor. The voltage waveforms of a back-to-back energizing event look very much the same as those in normal energizing. Almost all features of normal energizing apply to back-

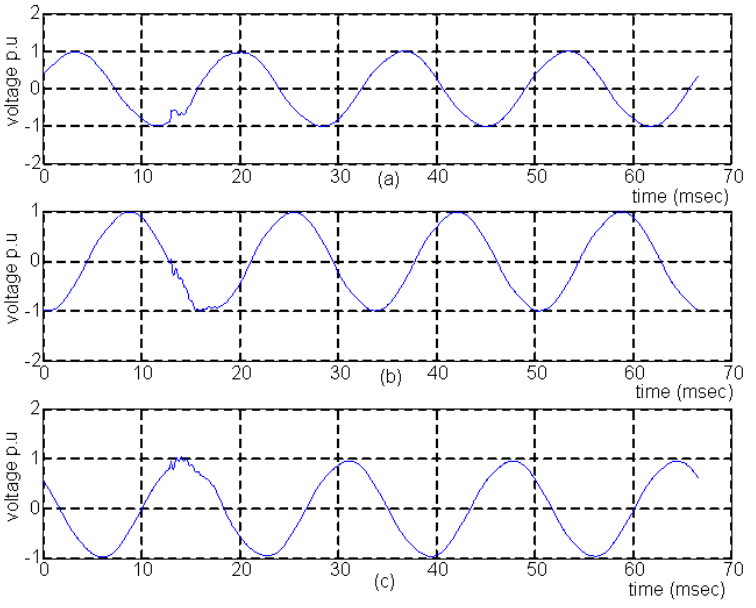


Fig. 9. Typical voltage waveforms of back-to-back capacitor energizing of utility capacitors in phases (a) A, (b) B, and (c) C

to-back energizing as well. Therefore, discriminating between normal and back-to-back energizing events is a very delicate task. Fig. 9 shows a three-phase voltage of back-to-back energizing of a utility capacitor bank. There are, however, two key features that may help distinguish back-to-back from normal energizing-the step voltage change and the oscillatory frequency of the phase current.

- **Step Voltage:** Back-to-back energizing transients involve two capacitors in close vicinity. One of them is fully energized and then the other is switched on. The resulting step voltage is not as large as it would be if the fully energized bank was absent. However, this feature may not always be used to differentiate back-to-back from normal energizing. The step voltage in back-to-back energizing behaves exactly like the step voltage observed when the power quality monitor is located far away from the capacitor during normal energizing. Based on our observations, the voltage change in back-to-back capacitor energizing is typically no larger than 0.5 p.u.
- **Oscillation Frequency of the Energizing Event at the Phase Current:** The frequency of the phase current provides distinct evidence for the back-to-back energizing event. The typical frequency for back-to-back energizing in the phase current is greater than 1000 Hz, whereas that for the normal energizing is less than 1000 Hz. Fig. 10 shows the energizing transient in the phase current. The frequency in these phase currents is obviously higher than those under normal energizing (see Fig. 8).

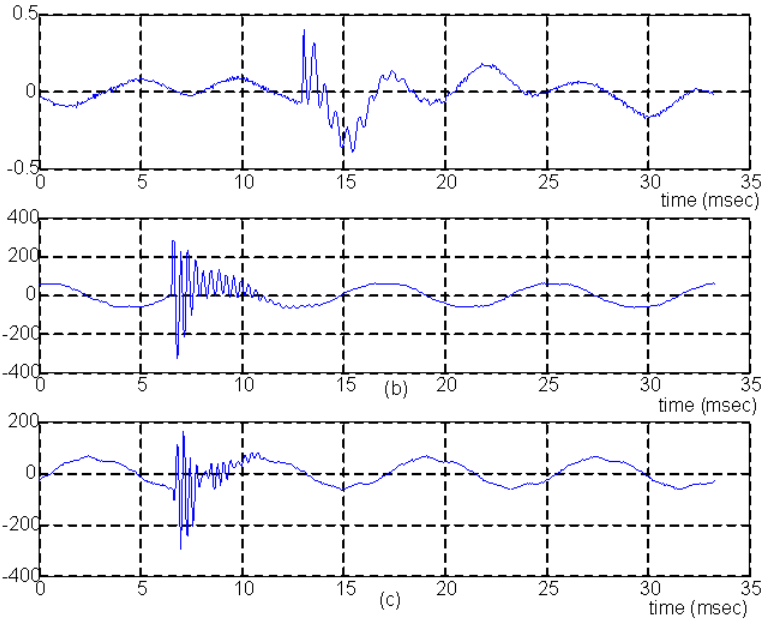


Fig. 10. Current waveforms at phases A, B, and C of the back-to-back capacitor energizing of utility capacitors

2.2 Classification of PQ Events

Recent advances in computer-based measurement technologies have allowed engineers to deploy intelligent systems[7] in PQ monitoring and analysis devices. The primary goal of intelligent system applications in PQ measurement involves automatic classification of PQ disturbances in terms of their types and root causes [5]. Thus, the intelligent systems help provide critical information and prioritize recommended actions to manage critical equipment or avoid system failure.

It may be beneficial to compare two representative AI techniques – i.e., artificial neural network (ANN)-based intelligent systems and rule-based expert systems – that have been widely applied in PQ classifications.

2.2.1 ANN-Based Classifier

The artificial neural network (ANN), often called the neural network, is regarded as the most generic form of AI for emulating the human thinking process. An ANN tends to emulate the biological nervous system of the human brain in a very

[7] Artificial intelligence (AI) techniques such as expert systems (ES), fuzzy logic (FL), artificial neural networks (ANN), genetic algorithms (GA), and synergistic hybrid schemes have been applied extensively to solve many industrial problems, including problems in process control, healthcare, geology, agriculture, information management, homeland security, and space technology, as well as power systems, to name a few [16].

limited way using computer programming or electronic circuitry. An ANN consists of the interconnection of a large number of nonlinear processing units referred to as neurons. The internal structure of the ANN may involve feed-forward paths only, or both feed-forward and feedback paths. The most important properties of the ANN approach are summarized below [16]:

- **Nonlinearity.** This property is particularly useful if the underlying physical mechanism responsible for the generation of an input signal is inherently nonlinear.
- **Weak statistical assumptions.** ANN relies on the availability of training data for its design; it is therefore able to capture the statistical characteristics of the environment in which it operates, provided that the training data are sufficient to represent the environment.
- **Learning.** ANN has a built-in capability to learn from its environment by undergoing a training session for the purpose of adjusting its free parameters.

These properties have justified the use of a neural network approach for pattern-recognition and image-processing problems that are particularly difficult to solve. Because the issue of PQ disturbance waveform identification is essentially a problem of pattern recognition, many ANNs have been used to classify power-quality disturbance events and, shown promising results.

One of the most notable of these is waveform recognition using a wavelet-based neural classifier [8, 6], which performs waveform recognition in the wavelet domain using multiple neural networks. Wavelet transformation is very sensitive to signal irregularities but very insensitive to regular signal behavior. In other words, wavelet transform coefficients associated with the disturbance event have very large magnitudes while those of a disturbance-free waveform have small magnitudes.

Fig. 11 shows the conceptual block diagram of the wavelet-based neural classifier consisting of preprocessing, processing, and post-processing phases. In this case, the time domain PQ disturbance waveform is transformed into the wavelet domain before being fed to the neural network. Thus, the input of the neural network is a preprocessed signal. This transformation detects and extracts disturbance features in the form of time and frequency data. The extracted data help the neural network in distinguishing one disturbance event from another. The processing phase contains a set of multiple ANNs, which perform the waveform recognition via the wavelet domain (the domain of its input data). Because multiple neural networks are utilized, a post-processing phase is required to combine the outcomes of the multiple neural networks in order to make a decision about the disturbance type and to provide a level of confidence for the decision made.

One of the drawbacks of using ANN is that it can only identify PQ events for which it has been trained. If the incoming waveform is not one of the events for which the ANN has been trained, it will incorrectly identify the waveform. Another drawback of the ANN approach is its lack of scalability. The algorithm developer cannot adjust ANN structures (e.g., the number of input and output stages, hidden layers). Therefore, making changes to incorporate the identification of a new PQ event requires retraining of the whole ANN system. However, although its performance may be limited and it is not readily scalable, the ANN approach is a preferred

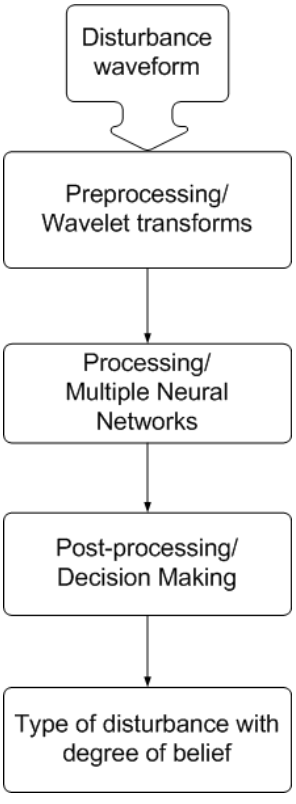


Fig. 11. Conceptual block diagram of the wavelet-based neural classifier

choice when data are abundant and PQ experts are not available. In order to overcome these disadvantages, we may consider a scalable rule-based expert system while incorporating a reinforced self-adapting capability. This system is scalable, easily modified, and capable of mimicking a PQ engineer's thought process in PQ event identification.

2.2.2 Rule-Based Expert System

Identification of PQ waveforms usually requires a significant amount of PQ expertise. A trained PQ engineer can easily identify the cause of a disturbance from voltage and/or current waveforms. Identifying a handful of disturbance waveforms is typically not problematic and can be performed with a high degree of accuracy. However, identifying thousands of disturbance waveforms at a time is certainly an onerous task. In addition to the drawback of human error, manual procedures used to identify power-quality events are obviously not practical and quite costly. It is therefore desirable that the expertise of power engineers be reproduced and coded into a set of programs. This set of computer programs will behave as if it were an

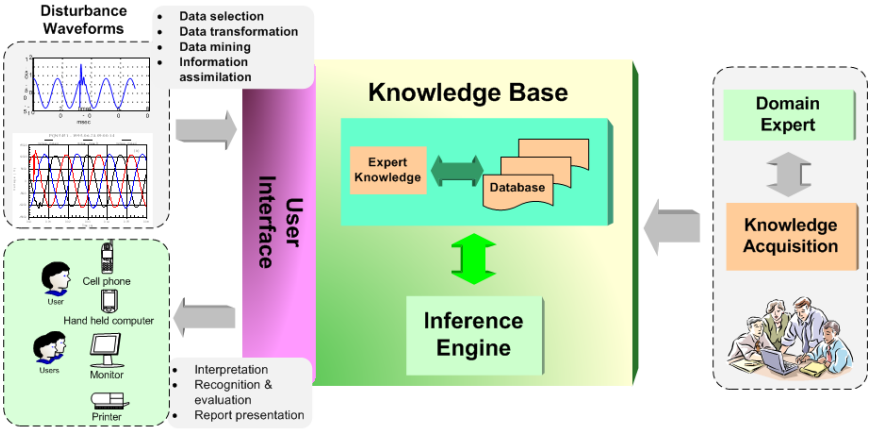


Fig. 12. Structure of a rule-based expert system

expert in power quality – it will possess the artificial intelligence necessary to perform various power-quality analysis tasks [7].

Fig. 12 shows the general structure of a rule-based expert system designed for a PQ application. The inputs of the expert system are voltage and/or current waveforms. The output of the expert system is the desired information or actionable intelligence. The expert system consists of three basic elements:

- **Inference engine or the control procedure mechanism.** This element draws inferences based on previously available knowledge and controls the analytical flow.
- **Knowledge base.** This is a compendium of static knowledge represented by production rules (or if-then rules). For this reason, an expert system is often called a rule-based system.
- **User interface.** This element facilitates communication between users and the expert system.

The structure of the expert system described above is indeed a common structure found in many expert systems. The flowchart for application program development is shown in Fig. 13, which is similar to Fig. 11 except that here the system structure is represented in a more developmental form and the if-then rules are delineated.

In a large and complex knowledge base, meta-rules and other forms of meta-knowledge can increase the efficiency with which the expert system reaches a conclusion. Unlike a conventional program, an expert system is said to have a learning capability because of meta-knowledge. Meta-knowledge is knowledge about the operation of a knowledge base, and meta-rules are rules about the operation of the rules. Meta-knowledge determines the most efficient operation strategy and allows the knowledge base to learn which rules are most useful from experience. Avoidance of testing for unlikely rules will improve the efficiency of a knowledge base search. This is particularly true with an expert-system-based real-time control and

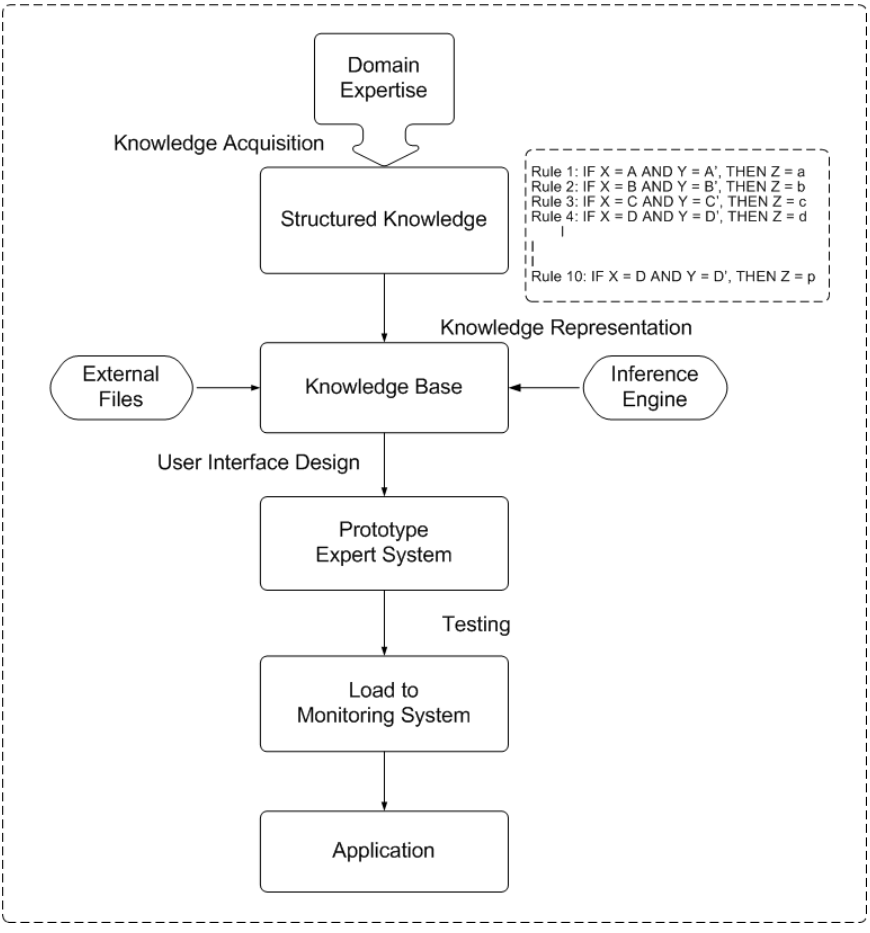


Fig. 13. Expert system program development flowchart

monitoring of power systems. An example of a meta-rule to guide the order of a rule search within a given frame is:

Meta-Rule 1: IF rules 2, 5, and 9 are skipped 15 times consecutively AND rule 1 is never tested THEN do not test rule 1 AND test rules 2, 5, and 9 at the end.

A successful expert system in PQ can be constructed if the following prerequisites are available:

- There must be at least one human expert to identify and interpret disturbance waveforms under consideration. The human expert is the source of knowledge in the power-quality domain. In an ideal situation, the human expert is a group of individuals with thorough knowledge in power quality.
- There must be at least one engineer to formalize and encode the knowledge of the human expert into a set of computer programs. In addition, an expert in extracting

power-quality signatures or empirical features from disturbance waveforms is needed. Extracting power-quality signatures requires in-depth signal-processing techniques.
- There must be a set of disturbance waveforms or data to provide empirical features and to validate computer codes that embody the knowledge.

Let's revisit Fig. 3, which shows the building blocks of the scalable expert system for the automated evaluation of capacitor-switching transient concerns [18]. There are six knowledge bases that are essentially autonomous software modules. These modules perform the most essential tasks for evaluating capacitor-switching transients, and together they complete the automated system.

Each knowledge-base module stores knowledge (in the form of if-then rules) relevant to each specific task. An individual knowledge base extracts predefined characteristic disturbance features from a given disturbance waveform, and these features are then compared with the stored knowledge. If a match is made, then a PQ event has been identified, otherwise no match is made and the event is classified as unidentified. The outcome of the comparison is then sent to an inference/procedure control engine. The waveform, along with the extracted features, is now passed on to the next module. The order of analysis depends on the author(s) of the expert system. However, the module data flow is usually ordered according to increasing identification difficulty (i.e., an increasing number of features to be extracted).

Unlike ANN approaches to identifying PQ events, this approach is easier to modify, debug, and expand. In ANNs, a weight factor connecting one neuron to another is fixed – one cannot change its value without ANN retraining. This issue also makes adding new PQ events to an already-trained ANN troublesome. Furthermore, if the outputs of the ANNs do not give a correct PQ event type, it is not possible to debug the ANN other than to retrain the whole network with more data or to improve preprocessing techniques.

In summary, the rule-based PQ evaluation system possesses the following merits:

- The rule-based system is easy to develop and maintain. The system is capable of reproducing a PQ engineer's thought process in identifying the six capacitor-switching transient characteristics specified above. Should identification of a new capacitor-switching transient characteristic be desired, a new knowledge base module can be created and added to the scalable system. Thereby, the capability of the system can be expanded at will.
- Rules in the scalable system are also easy to modify and debug. If a disturbance event causes magnification of capacitor-switching transients, but the output reports that the event does not cause a magnification issue, then it is understood that one or more rules in the module are not appropriate or that perhaps more rules are needed.
- It has the ability to declare a given waveform as unidentified. For example, a waveform with a capacitor switching transient event of category X is entered to the system. It begins analyzing the waveform through all modules. The output of each module is certainly no match. Thus, after all possibilities are exhausted, the waveform is declared as unidentified, thereby minimizing misclassification errors.

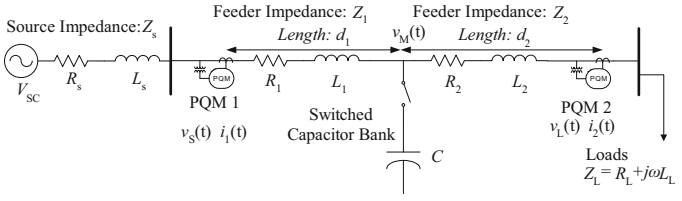


Fig. 14. The switched capacitor bank is downline from power-quality monitor PQM 1 and upline from PQM 2. The CT direction is in the direction of feeder current flow.

2.3 Correlation of PQ Events with the System Equipment

As discussed previously, it should be beneficial to correlate the PQ disturbance with the relevant system equipment that facilitates the disturbance. For example, overvoltage transients resulting from capacitor switching operations can cause adverse impacts both on the power system and within customer facilities. Locating the problematic capacitor bank helps the utilities prevent any PQ problems or mitigate the impact by taking timely measures [22].

The problem addressed in this section can be stated as follows: Given voltage and current waveforms resulting from the capacitor-energizing event, determine the relative direction of the switched capacitor bank. That is, let us consider a one-line diagram illustrated in Fig. 14, where a capacitor bank is energized. A power-quality monitoring device (PQM) is located on both sides of the capacitor bank with its current transformer (CT) direction[8] pointing to the direction of the feeder current. In this arrangement, the switched capacitor bank is physically downstream of PQM 1, but upstream of PQM 2. Thus, we determine only the physical location of the switched capacitor bank in order to be consistent in presenting the direction.

Several techniques have been proposed to determine the location of switched capacitors. Parsons [24] investigated the disturbance energy flow during the transient period and the polarity of the initial peak of the disturbance power to find out the relative location of the switched capacitor bank. The technique requires a provision that the disturbance energy must be greater than or equal to a certain percentage of the peak excursion of the disturbance energy. If this condition is satisfied, the technique subsequently compares polarities of final values of disturbance energy and power. Kim [34] applied a backward Kalman filter to find the location of a switched capacitor bank by estimating the voltage rise of the capacitor bank. Chang [23] proposed a method using voltage-disturbance energy index and current branch current phase-angle variations for tracking transients. This section presents in more detail research based on the two fundamental physical phenomena of capacitor bank energizing.

[8] The capacitor bank can be considered to be electrically upstream of PQM 1 but downstream of PQM 2 when both CTs are reversed upon installation. Nevertheless, its physical location is unchanged. When the CT is found to be installed reversely, the measured current could be flipped in the preprocessing stage.

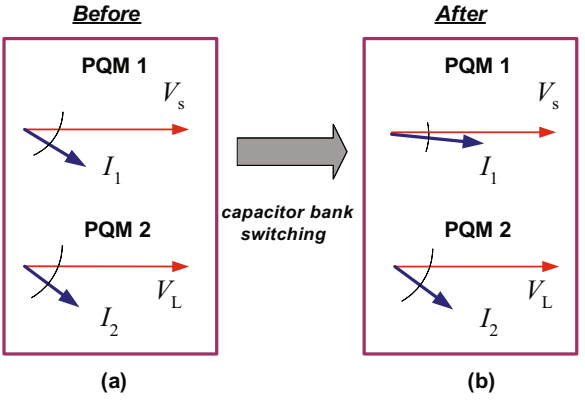


Fig. 15. Change in the phase-angle relationship between voltage and current waveforms before and after the capacitor bank energizing. Only PQM1 can detect the phase-angle jump.

The underlying physical properties are as follows. First, reactive power flow only upstream of the capacitor bank is reduced by the injected reactive power. However, reactive power flow downstream of the capacitor bank is not affected at all. Thus, we will observe the power factor improvement only at the upstream PQ monitors, as illustrated in Fig. 15. Second, when an uncharged capacitor bank is energized, it behaves like an instantaneous short-circuit element because the initial voltage at the capacitor terminal is zero. As a result, the instantaneous system voltage at the energizing instant is pulled towards zero voltage. If a measurement is available at the source-side of the capacitor bus, the instantaneous system voltage would be zero. Due to the instantaneous short-circuit phenomenon, current measured at the substation (or upstream from the capacitor location) will have a significantly large magnitude but short-time current magnitude (i.e., inrush capacitor currents). On the other hand, current measured downstream from the capacitor sees little of the inrush current. As the capacitor is being charged, the instantaneous voltage rises and overshoots its system peak voltage momentarily. Thus, the relative location of a switched capacitor bank can be determined by observing the polarities of initial voltage and current change – i.e., the signs of voltage gradient, dv/dt, and current gradient, di/dt, at the monitoring locations. Fig. 16 illustrates these phenomena clearly. The accuracy and efficacy of the direction-finding technique above can be demonstrated analytically [37], as well as by way of time-domain simulation models and actual data [35, 22].

In order to utilize these two underlying physics, accurate detection of the switching event is required. Algorithms to calculate the time derivatives of voltage and current immediately after the switching event are also required. By comparing the zero crossings of voltage and current waveforms based on the time stamps, one can also calculate the phase-angle changes before and after the switching event. Cross time-frequency [13] analysis can also provide time- and frequency-localized phase difference between the transient voltage and current disturbance waveforms.

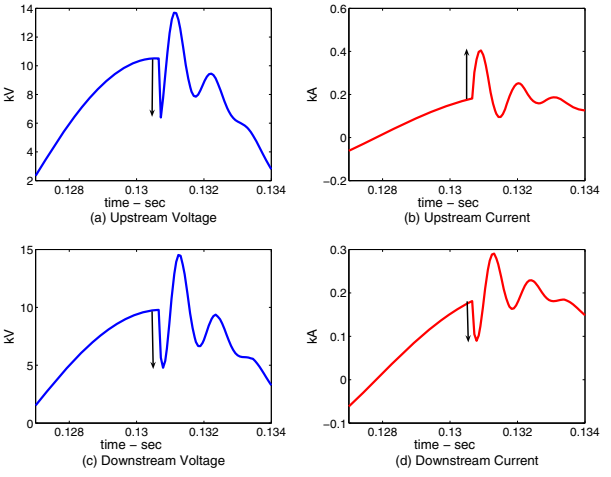


Fig. 16. Voltage and current waveforms recorded from power-quality monitors, which are upstream (a and b), and downstream (c and d) from the switched capacitor bank. Note that initial changes of current on b and d are dissimilar.

2.4 Modal Properties of Electric Power Systems

The modal properties of the electric power systems provide very useful information to help develop any preventive control actions or execute mitigation schemes against system resonance problems. As we discussed earlier, harmonic resonance in a utility distribution system can occur when the system natural resonant frequency is excited by relatively small harmonic currents from nonlinear loads [2]. The system voltage and current may be amplified and highly distorted during this resonance encounter. This scenario is more likely to occur while the energizing of a capacitor bank in a system has little or negligible resistive damping. Thus, it is desirable to predict the likelihood of harmonic resonance using system damping parameters such as the Q factor[9] and the damping ratio ζ at the resonance frequency. Various signal-processing techniques help extract this damping and resonant frequency information from the PQ measurement data, especially from the transient data such as capacitor bank transient data. Accurate system modal information can be utilized to evaluate the system's vulnerability to power quality disturbances, particularly resonance phenomena, so that a utility can take preventive measures and improve the PQ of the system.

[9] The Q factor is more commonly known as the X/R ratio. The reactance and resistance forming the Q factor should be the impedance effective values that include the effect of loads and feeder lines, in addition to impedances from the equivalent Thevenin source and substation transformer. In other words, the X/R ratio is influenced by the load level. When the ratio is high, harmonic resonance is more likely to occur.

2.4.1 Estimation of Resonant Frequency

A simple yet accurate empirical method using FFT can estimate the system parallel resonant frequencies using voltage and current waveforms (less than a cycle) from capacitor bank energizing events, and their FFT analysis results, leading to the spectral analysis of the system impedance as follows [39, 19]:

$$Z(f) = \frac{FFT\,of\,V(t)}{FFT\,of\,I(t)} = \frac{V(f)}{I(f)}.$$

Fig. 17 illustrates a real-world example of this method and clearly indicates the resonant frequencies at around 500 Hz and 720 Hz.

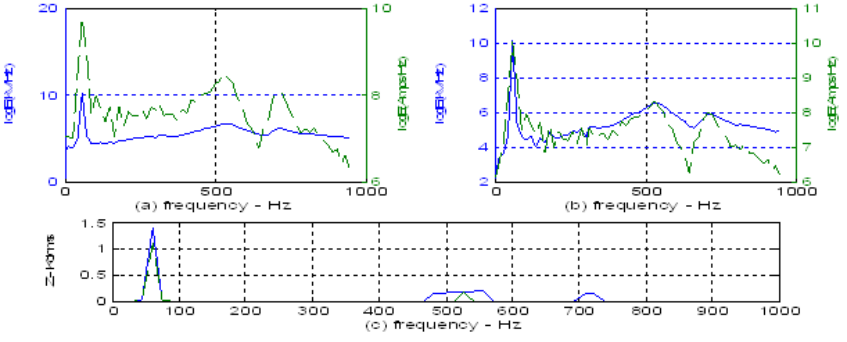


Fig. 17. Resonant frequency estimates: (a) FFT of voltage (b) FFT of current and (c) FFT of impedance

An improved method using only the intrinsic transient portion of the capacitor bank transient was also proposed and was proven to effectively remove the artificial frequency components that may mislead the users [20]. The underlying idea is that the capacitor bank transient will be composed of steady-state data and intrinsic transient data, which will be damped out within a short time. Thus, the intrinsic portion can be extracted by removing the steady data from the whole transient data. Of course, accurate detection of the switching event is required for these two methods [8].

2.4.2 Estimation of Damping Parameters

A few studies on the application of signal processing techniques to transient modeling and analysis have been undertaken on the assumption that transient components are exponentially damped sinusoids. These techniques include Prony analysis and Hilbert transform-based or wavelet transform-based system-identification techniques, etc. [27, 25, 26]. ESPRIT and other system-identification techniques such as the All-Pole model have also been applied [6]. However, they will not be discussed in this chapter.

Prony Analysis

Given accurate system parameters, eigenvalues from linear matrices arising from the dynamic model of a system may provide analytical modal parameters. However, the actual system may be too complex to model or may have time-varying parameters. In these cases, it may be practical to substitute an estimated linear model for the actual dynamic model. The estimated model can be derived from the system output waveform. One of viable approaches to estimate the system parameters is Prony analysis [27, 28, 29].

Prony analysis is basically designed to fit the parameters of a sum of damped sinusoidal functions to an observed measurement.

$$\hat{v}(t) = \sum_{i=1}^{n} a_i e^{-\zeta_i \omega_{ni} t} \cos\left(\omega_{di} t + \phi_i\right) = \sum_{i=1}^{n} \alpha_i e^{\lambda_i t} \tag{1}$$

where $v(t)$ consist of N samples (i.e., $v(t_k) = v(k), k = 0, 1, \ldots, N-1$) that are evenly sampled by Δt.

Fig. 18 illustrates how closely the reconstructed waveform through the Prony analysis can fit in the real measurement data from 115 kV line.

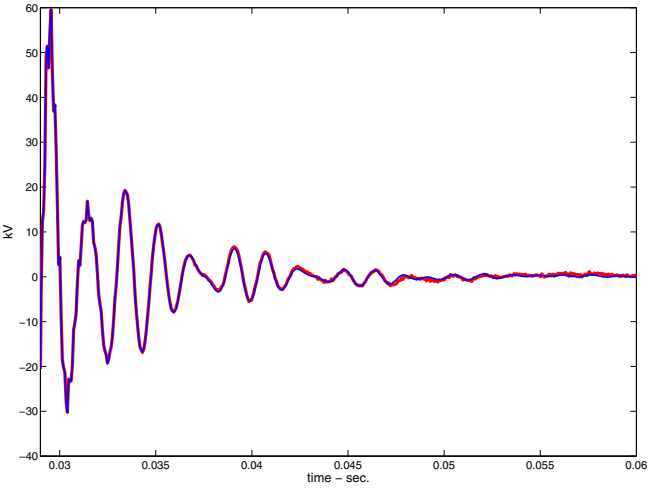


Fig. 18. Original voltage transient (bold line) and Prony estimate (solid line) from 115 kV line

In practice, the measurement may be corrupted by high-frequency noise or offset by a dc value, and thus one may need to preprocess the measurement before applying the Prony method. In general, the Prony analysis requires a computationally complex series of procedures for building a linear prediction model, then calculating the roots of the characteristic equations and amplitude and phase for each mode. Thus, it is often criticized for its computational intensity and artificial modal

information generated to fit in the waveform. The appropriate system order that should be determined beforehand is not generally known, and it may require a trial-and-error process.

Hilbert Transform-based Damping Analysis

The Hilbert transform can be applied to determine and extract the circuit properties embedded in the envelope of the waveshape of the transient signal such as capacitor switching transient [25]. The Hilbert transform of a real-valued time domain signal $y(t)$ is another real-valued time domain signal, $\tilde{y}(t)$, such that an analytic signal $z(t) = y(t) + j\tilde{y}(t)$ exists [36]. This is a generalization of Euler's formula in the form of the complex analytic signal. It is also defined as a 90-degree phase-shift system, as shown below:

$$\tilde{y}(t) = H[y(t)] = \int_{-\infty}^{\infty} \frac{y(\tau)}{\pi(t-\tau)} d\tau = y(t) * \left(1/\pi t\right) \tag{2}$$

$$F[\tilde{y}(t)] = \tilde{Y}(f) = (-j\text{sgn}f)Y(f) \tag{3}$$

where $\tilde{Y}(f)$ is the Fourier transform of $\tilde{y}(t)$. From $z(t)$, we can also write $z(t) = a(t) \cdot e^{j\omega(t)}$, where $a(t)$ is the envelope signal of $y(t)$ and $\theta(t)$ is the instantaneous phase signal of $y(t)$. The envelope signal is given by $a(t) = \sqrt{y(t)^2 + \tilde{y}(t)^2}$ and the instantaneous phase, $\theta(t) = \tan^{-1}\left(\frac{\tilde{y}(t)}{y(t)}\right)$. Using the property of Eq. (3), one can easily obtain the Hilbert transform of a signal, $y(t)$. If we let $Z(f)$ be the Fourier transform of $z(t)$, we can obtain the following relationship:

$$z(t) = F^{-1}[Z(f)] = y(t) + j\tilde{y}(t) \tag{4}$$

Thus, the inverse Fourier transform of $Z(f)$ gives $z(t)$, as shown in Eq. (4).

$$y(t) = y_m e^{-\zeta\omega_n t}\cos(\omega_d t + \phi) \tag{5}$$

$$\tilde{y}(t) = y_m e^{-\zeta\omega_n t}\sin(\omega_d t + \phi) \tag{6}$$

For the case of quadratic damping, the decaying transient and its Hilbert transform can be represented as Eqs. (5, 6). Thus, the resulting envelope, $a(t)$, becomes $y_m e^{-\zeta\omega_n t}$, where y_m is an arbitrary constant magnitude. This is a unique property of the Hilbert transform applicable to envelope detection.

In fact, the envelope from the Hilbert transform is not an ideal exponential function and is full of transients especially for those low-magnitude portions of the signal approaching the steady-state value (ideally zero). Thus, only a small number of data are utilized in order to depict the exponential satisfactorily: one cycle of data from the capacitor-switching instant is generally sufficient to produce a good exponential shape. The number of data elements will depend on the sampling rate of the PQ monitoring devices and should be calibrated by investigating the general load condition, especially when the method is applied to a new power system in order to optimize the performance. The obtained data are now fitted into an exponential

function. The direct way to fit the data into the exponential function is possible through the iteration-based nonlinear optimization technique. However, the exponential function is namely an intrinsic linear function, such that the $\ln a(t)$ produces a linear function; i.e.,

$$\ln\{a(t)\} = \ln y_m - \zeta\omega_n t. \tag{7}$$

As a result, we can apply the standard least-squares method to approximate the optimal parameters more efficiently [28]. The solution is not optimal in minimizing the squared error measure, due to the logarithmic transformation. However, except for very high damping cases, this transformation plus the least-squares estimation method, creates a very accurate estimate of $a(t)$.

Analytic Wavelet Transform-based Damping Analysis

The main problem of the Hilbert damping analysis is that it does not perform as well for multi-modal cases[10] as for uni-modal cases. Unfortunately, it is not uncommon to find many practical cases for which multiple modal components are comparable in size such that there is no conspicuous single mode to study. One may reinforce the signal-preprocessing stages by adding high-frequency noise-rejection filters and/or adding band-pass filters and thus appropriately selecting the important resonant frequencies based on the system studies followed by the Hilbert damping analysis.

It is also interesting to note that the continuous wavelet transform can be applied to identify the dynamic system modal parameters. These research problems have been encountered in civil and mechanical engineering and have shown very promising results [40]. The vibration data, acoustic data, and ambient noise have been utilized to identify system modal parameters, from which one can detect any damage of the system in order to prevent further damage. It was demonstrated that the wavelet-based system-identification method can be applied to the capacitor bank transient data to characterize electric power systems [26].

An analytic wavelet can be constructed by the frequency modulation of a real and symmetric window $g(t)$ – i.e., $\psi(t) = g(t)\exp(j\eta t)$ [38]. The Fourier transform of $\psi(t)$ is then $\hat{\psi}(\omega) = \hat{g}(\omega - \eta)$. Thus, $\hat{\psi}(\omega) = 0$ for $\omega < 0$ if $\hat{g}(\omega) = 0$ for $|\omega| > \eta$. Hence, ψ is analytic. Specifically, a Gabor wavelet is a representative of an analytic wavelet transform and is obtained from a Gaussian window [38]:

$$g(t) = \frac{1}{(\sigma^2\pi)^{1/4}}\exp\left(\frac{-t^2}{2\sigma^2}\right) \tag{8}$$

The Fourier transform of this window is then computed as follows:

$$\hat{g}(\omega) = \left(4\sigma^2\pi\right)^{1/4}\exp\left(\frac{-\sigma^2\omega^2}{2}\right) \tag{9}$$

[10] Two or more energized shunt capacitor banks and feeders may interact and form two or more modal (resonant) components. These cases correspond to multi-degree-of-freedom (MDOF) systems in mechanical or structural engineering [40].

Thus, if $\sigma^2\eta^2 \gg 1$, $\hat{g}(\omega) \approx 0$ for $|\omega| > \eta$. Hence, Gabor wavelets are considered to be analytic. The Fourier transform of $\psi_{u,s}$ is a dilation of $\hat{\psi}(\omega)$ by $1/s$, and can be obtained as $\hat{\psi}_{u,s}(\omega) = \sqrt{s}\hat{\psi}(s\omega)e^{-j\omega u}$. Thus, the AWT of $f(t)$ (i.e., $Wf(u,s)$) is the inverse Fourier transform of the frequency function obtained by multiplying $\hat{f}(\omega)$ by $\hat{\psi}(s\omega)$. Thus, given the Fourier transform of the chosen analytic wavelet transform, computational complexity can be reduced. This is because the Fourier and inverse Fourier transforms of the signal, $f(t)$ at each scale, s, can be calculated efficiently via FFT[11] [39]. In practice, this is a significant benefit of the proposed method because it can be easily implemented in the existing monitoring devices.

An analytic wavelet transform defines a local time-frequency energy density named a scalogram:

$$P_W f(u,s) = |Wf(u,s)|^2 \tag{10}$$

Note that we focus on the instantaneous frequency, $\omega(u)$, defined as $\eta/s(u)$. This is related to the ridges[12] of the normalized scalogram, $s^{-1}P_W f(u,s)$ [38].

Let us represent a signal, $v(t)$, as $v(t) = a(t)\cos\theta(t)$. This may represent an entire single-mode signal or a single-modal component of a multi-modal signal. The Gabor wavelet transform of this signal is obtained as follows [40]:

$$Wv(u,s) = \frac{1}{2}a(u)\,\hat{\psi}_{u,s}(\omega(u),\sigma,\eta)\,e^{j\theta(u)} + \varepsilon\left(a'(t),\omega'\right), \tag{11}$$

where $\omega = \theta'(u) = \frac{d\theta(u)}{du}$, $\omega'(u) = \frac{d\omega(u)}{du}$, and $\hat{\psi}_{u,s}$ is a dilated and translated version of (9),

$$\hat{\psi}_{u,s}(\omega,\sigma,\eta) = \left(4\sigma^2 s^2\pi\right)^{1/4} e^{\left(\frac{-(\omega-\eta/s)^2\sigma^2 s^2}{2}\right)} e^{-j\omega u}. \tag{12}$$

The last term, $\varepsilon(a'(t),\theta''(u))$ in Eq. (11) is an approximation error and can be neglected if the derivative of the phase is greater than the bandwidth $\Delta\omega$ [38]; i.e., $\theta'(u) \geq \Delta\omega$. The derivation for this Gabor wavelet function is presented in [40].

In general, the signal $v(t)$ can be represented as a damped sinusoidal: $v(t) = \alpha e^{-\zeta\omega_n t}\cos(\omega_d t + \phi)$, where $\omega_d = \omega_n\sqrt{(1-\zeta^2)}$ and ω_n is a system natural frequency. The algorithm for estimating the power-system damping ratios based on AWT can be developed as follows:

1. The signal utilized for estimating the system damping is the same capacitor bank switching transient as used in [20, 26], where an empirical method was taken to

[11] The discrete Fourier transform of a circular convolution is the product of the two discrete Fourier transforms. For two signals f and h with a length $M \geq 32$, computing their convolution with an FFT is faster than using the straightforward formula. In this case, it is required to calculate $\hat{g}(\omega)$ and $\hat{f}(\omega)$ only once. It is then required to multiply these functions by varying scales. The inverse FFT is finally required only once. Note that the spectral range of interest will determine the number of required scales.

[12] The time-scale representation of the energy concentration of the continuous wavelet transform is called the ridge. Ridges represent the frequency content of the analyzing signal with a high density of energy that is a function of u [38].

extract the free response of the system to the capacitor bank energizing. This free response is dictated by the system resonant frequencies.
2. Perform FFT on this free response of the system and determine the salient ω_d's. Note that ω_d's replace ω_n's in this algorithm development, since the observable resonant frequency components are ω_d's. However, small damping ratios make the numerical difference between them negligible.
3. Obtain the Gabor wavelet coefficients using Eqs. (11) and (12).
4. The angular velocity of the Gabor wavelet function ($\omega = \eta/s$) is equal to the angular velocity of the signal i.e., ω_d as far as the ridge is concerned. Therefore, $\exp\left(-\left(\omega - \eta/s\right)^2 \sigma^2 s^2 \Big/ 2\right)$ will be 1.
5. Consequently, the following relation can be derived:

$$\ln\left(\frac{2\left|Wv\left(u, s(u)\right)\right|}{\left(4\pi\sigma^2 s(u)^2\right)^{1/4}}\right) \approx -\zeta\omega_d u + \ln\alpha. \tag{13}$$

It should be noted that Eq. (13) is a linear function of u. Thus, the standard linear squares method can be applied to find the slope of Eq. (13). As to a multi-modal signal, this linear regression analysis should be performed for all modal components of interest. The ridge is considered to be constant because a constant, instantaneous angular velocity is assumed. Since $s(u) = \frac{\eta}{\omega_d}$, where η denotes the frequency center of the base wavelet, the frequency center of a dilated wavelet is η/s.

It is necessary to tune the parameters (η and σ) of the Gabor wavelet to obtain accurate damping ratios. This process of tuning can be systematically conducted based on the the bandpass filter design methodology [39] and empirical observations of the capacitor-switching transients [26].

In designing the bandpass filter – i.e., selecting the parameters of the Gabor wavelet – one should consider the center frequency, bandwidth[13] and the associated quality factor, Q. In our case, the center or resonant frequency is already known (estimated). Bandwidth or Q value can be calculated once either of them is determined.

A synthetic voltage signal, $v(t)$ is composed of three damped sinusoidal components and can be represented as follows:

$$v(t) = \sum_{i=1}^{3} \alpha_i e^{-\zeta_i \omega_{ni} t} \cos\left(\omega_{di} t + \phi_i\right). \tag{14}$$

This signal is composed of the 5th, 7th, and 11th harmonics (modes) with reference to the system frequency (60 Hz). The magnitudes of modal components are selected based on our experience in harmonic analysis with reference to the magnitude of the 5th harmonic voltage, α_1. Thus, $\alpha_1 = 100$ (selected arbitrarily), $\alpha_2 = 100 \cdot 5/7 = 71.43$, and $\alpha_3 = 100 \cdot 5/11 = 45.45$. The damping ratios, ζ_i's, for these harmonic components are randomly selected as 0.03, 0.05, and 0.1, respectively. The phase angles, ϕ_i's, are chosen arbitrarily as $\pi/3$, $-\pi/4$, and $\pi/6$,

[13] It is defined to be the frequency range between the -3dB points located on either side of the center frequency.

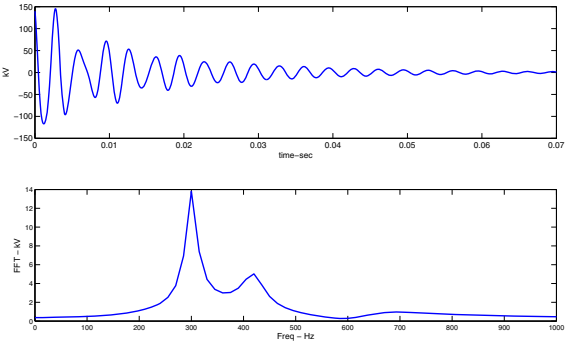


Fig. 19. (a) Synthetic voltage transient waveform, and (b) FFT of the waveform

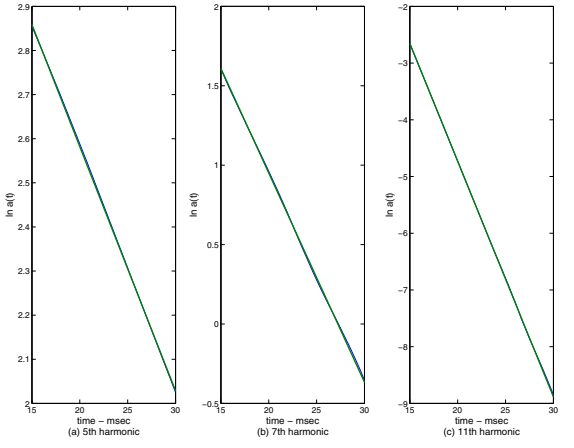


Fig. 20. Linear regression analysis for (a) 5th harmonic, (b) 7th harmonic, and (c)11th harmonic

respectively. Fig. 19 shows (a) the synthetic waveform and (b) the FFT result for this signal. Fig. 20 illustrates the linear regression analysis based on Eq. (13) for (a) the 5th mode, (b)the 7th mode, and (c) the 11th mode of the signal. Note that the solid lines are fitted into the bold lines. Based on (13), linear regression analysis finds the slopes of each mode, $(-\zeta_i \omega_{di})$.

Using ω_{di}'s obtained from the spectral analysis of the signal, we can obtain the multi-modal damping ratios $(\zeta_i' s)$. Based on the assumption that the voltage in the s-domain can be represented by a sum of quadratic equations, we can interpret $(\frac{1}{2\zeta_i})$ as the effective X/R ratio at ω_{di} as defined in [26].

As described in Table 1 above, the estimated damping ratios are very close to the theoretical values.

Table 1. Estimation Results for a Synthetically Composed Signal

Parameters	$f_d = \omega_d/2\pi$	ζ
Theoretical values	300, 420, 660	0.030, 0.050, 0.100
Estimates using AWT	303, 420, 694	0.030, 0.050, 0.095

2.5 PQ Data Management

Present data-reduction techniques in power-disturbance monitoring use a wrap around scheme that recycles the existing memory capacity by writing over the oldest data, and a quantization technique that represents a group of similar waveforms as a single entity. Both methods conserve memory space[14] but do not compress data in the usual sense of data compression [30, 31, 32].

Wavelet transforms can be utilized in compressing power-quality disturbance signals. Wavelet-based compression schemes have been utilized especially in the image-processing area. Wavelet transforms decompose a given signal into several scales at different levels of resolution. At each scale, the wavelet transform coefficients that correspond to a particular disturbance event are exclusively larger than those that do not correspond to the event in question. Therefore, related coefficients are kept along with their temporal information, while others that are not related to disturbance events are discarded. In this way, the amount of stored data can be reduced. To recover the original signal, reconstruction is performed using the most-smoothed signal, along with the saved wavelet transform coefficients of the detail signals. The compression scheme is similar to the noise reduction or denoising procedures described in [31], where a dyadic orthonormal wavelet was used in power-quality data compression techniques. With this transform, perfect reconstruction is guaranteed. In other words, the reconstructed signal is the same as the original signal as indicated in Eq.s (15) and (16) below.

$$DWT_\psi x(m,n) = 2^{-m/2} \int_{-\infty}^{\infty} x(t)\psi^*\left(\frac{t-n2^m}{2^m}\right) dt \tag{15}$$

$$x(t) = 2^{-m/2} \sum_m \sum_n DWT_\psi x(m,n)\psi\left(\frac{t-n2^m}{2^m}\right), \tag{16}$$

where $x(t)$ is the time domain signal to be decomposed or analyzed and Eq. (15) is the dyadic wavelet transform of $x(t)$; the asterisk denotes a complex conjugate;

[14] It is interesting to note that a disadvantage of using a high fixed sampling rate in the PQ monitoring is the large amount of data that must be stored in the monitor's physical memory for each recorded power-quality event. Although this issue is becoming less important today due to modern high-capacity memory devices a tradeoff is still made between the large memory requirements of high fixed sampling rates that enable faithful reproduction of the pertinent high frequency transient data versus lower sample rates that require less memory but may not faithfully capture all aspects of the waveform [17].

m and n are scale and time-shift parameters, respectively; and $\psi(t)$ is a given basis function (mother wavelet). Because the family of the dilated wavelets constitutes an orthonormal basis for $L^2(R)$, it is then possible to perfectly recover the original signal $x(t)$ from its coefficients, as presented in Eq. (16). With regard to wavelet data compression, the more scale decomposition one carries out, the more information is likely lost. However, this is not due to the dyadic orthonormal wavelet transform, but rather to the application of the threshold[15] the purpose of which is to discard undesired information – i.e., noise. However, in doing so, useful information may also be discarded. Therefore, the choice of threshold is very important. It was also observed that the compression ratio, defined as the original file size divided by the compressed file size, is in the range of 3–6 with normalized mean square errors of the order of 10^{-6} to 10^{-5}. Using the data-compression method, one can minimize data storage requirements and transmission time while preserving the reconstructed signal in such a way that it is virtually indistinguishable from the original.

A computationally simple but efficient data-compression method used in some instruments is a variable-sample-rate algorithm based on the rate-of-change of the waveform. It works on the principle that most of the time a lower sample rate will suffice and that a higher rate is only needed during each short burst of transient activity. This technique offers an effective means to capture high-frequency data while also limiting memory needs at times when high-frequency content is not present on the waveform. Not only does this allow the instrument with finite memory capabilities to store many more waveforms before it fills (allowing longer intervals between downloads), but also the smaller data files are less cumbersome to work with and analyze. Variable sampling-rate algorithms can significantly reduce memory requirements, depending on the nature of the data [17].

3 Summary and Future Directions

PQ monitoring has advanced significantly in recent decades not only because of more interest and need for it but also because of the much-improved instrument and analysis technologies. The modern time-synchronized, advanced microprocessor-based PQ recording devices can measure the full range of power conditions. The computational intelligence has also been enabling technologies for turning raw PQ measurements into a much more valuable commodity, i.e. actionable knowledge that can help us better understand PQ problems, evaluate situation, and correlate

[15] The choice of threshold varies according to the particular application. As pointed out by Donoho, the choice of threshold for denoising applications is based on the assumption that the noise in a given signal is white noise. From this assumption, the threshold level for each decomposition level is derived. However, for the application of compressing rms power system event data, the minimum length description approach as a threshold selection rule might be more appropriate. Threshold selection might be a good issue to investigate. For example, one might investigate the nature of the noise embedded in the power system and examine it in the wavelet in domain. A better threshold selection rule might then be derived from the statistical property of the noise in the wavelet domain.

the problems with the root causes, leading to the prevention and mitigation of PQ problems with timely controls. However, we cannot overemphasize the importance of performing PQ studies based on a working knowledge or understanding of the power systems in terms of the expected occurrence rates of typical disturbances, required measurement types and critical locations for them within the industry or customer expectations and guidelines. The PQ data should be analyzed within the context of facility load characteristics, utility system design, and the various applicable PQ guidelines because the type of power system affects how we interpret the PQ measurement data.

While this chapter covers the basic concepts of PQ monitoring and focuses on a few CI applications in PQ such as the capacitor bank switching transients, power system harmonics, and resonance, it conveys that as utilities and industrial customers have expanded their PQ monitoring systems, the data management, analysis, and interpretation functions have become the most significant challenges in the overall PQ research effort. It is highly desirable that analysis, characterization and classification of PQ disturbances, and generation of any useful PQ protection or mitigation schemes can be performed in an automated manner.

PQ monitoring is rapidly becoming an integral part of general distribution system monitoring, as well as an important customer service. Electric power utilities are integrating PQ and energy management monitoring, evaluation of protective device operation, and distribution automation functions. PQ information should ideally be available throughout the company via an intranet and should be made available to customers for evaluation of facility power-conditioning requirements. The PQ information should be analyzed and summarized in a form that can be used to prioritize system expenditures and to help customers understand the system performance. The information from PQ monitoring systems can help improve the efficiency of operating the system and the reliability of customer operations. These are benefits that cannot be ignored. The capabilities and applications for PQ monitors are continually evolving.

References

1. NETL: The Modern Grid Strategy: Principal Characteristics (2008), http://www.netl.doe.gov/moderngrid/characteristics.html
2. Dugan, R., McGranaghan, M., Santoso, S., Beaty, W.: Electrical Power Systems Quality. McGraw-Hill, New York (2003)
3. Santoso, S., Grady, W., Powers, E., Lamoree, J., Bhatt, S.: Characterization of distribution power quality events with Fourier and wavelet transforms. IEEE Trans. on Power Delivery 15(1), 247–254 (2000)
4. McGranaghan, M., Santoso, S.: Challenges and trends in analyses of electric power quality measurement data. EURASIP J. Adv. Signal Process. (Special Issue on Emerging Signal Processing Techniques for Power Quality Applications) (2007)
5. Bollen, M., Gu, I., Axelberg, P., Styvaktakis, E.: Classification of underlying causes of power quality disturbances using deterministic and statistical signal processing methods. EURASIP J. Adv. Signal Process. (Special Issue on Emerging Signal Processing Techniques for Power Quality Applications) (2007)

6. Bollen, M., Gu, I., Santoso, S., McGranaghan, M., Crossley, P., Ribeiro, M., Ribeiro: Bridge the gap bewteen signal and power. IEEE Signal Processing Magazine (July 12–31, 2009)
7. Santoso, S., Lamoree, J., Bingham, R.: Answermodule: autonomous expert systems for turning raw pq measurements into answers. In: International Conference on Harmonics and Quality of Power, vol. 2, pp. 499–503 (2000)
8. Santoso, S., Powers, E., Grady, W., Hoffman, P.: Power quality assessment via wavelet transform analysis. IEEE Trans. on Power Delivery 11, 924–930 (1996)
9. Gu, I., Bollen, M.: Time-frequency and time-scale domain analysis of voltage disturbances. IEEE Trans. on Power Delivery. 15(4), 1279–1284 (2000)
10. McCormick, S., Hur, K., Santoso, S., Maitra, A., Sundaram, A.: Capacitor bank predictive maintenance and problem identification using conventional power quality monitoring systems. In: IEEE PES General Meeting, Denver, USA, June 2004, pp. 1846–1850 (2004)
11. Lee, I., Dash, P.: S-transform-based intelligent system for classification of power quality disturbance signals. IEEE Trans. on Industrial Electronics 50(4), 800–805 (2003)
12. Zhao, F., Yang, R.: Power-quality disturbance recognition using S-Transform. IEEE Trans. on Power Delivery 22(2), 944–950 (2007)
13. Shin, Y., Powers, E., Grady, W., Arapostathis, A.: Power quality indices for transient disturbances. IEEE Trans. on Power Delivery 21(1), 253–261 (2006)
14. Santoso, S., Lamoree, J., McGranaghan, M.: Signature analysis to track capacitor switching performance. In: IEEE Power Engineering Society Transmission and Distribution Conf., pp. 259–263 (2001)
15. Bollen, M., Styvaktakis, E., Gu, I.: Categorization and analysis of power system transients. IEEE Trans. on Power Delivery 20(3), 2298–2306 (2005)
16. Jang, J., Sun, C., Mizutani, E.: Neuro-fuzzy and soft computing: A computational approach to learning and machine intelligence. Prentice Hall, New Jersey (1997)
17. Barker, P., Howe, B.: Power quality monitoring: Concepts, equipment, and applications, PQ Encyclopedia. In: EPRI (2009)
18. Santoso, S., McCormick, S., Hur, K.: Automated evaluation system for capacitor switching transient concerns. In: EPRI, Palo Alto, CA (2005), doi:1008510
19. Santoso, S., Maitra, A.: Empirical estimation of system parallel resonant frequencies using capacitor switching transient data. IEEE Trans. on Power Delivery 20(2), 1151–1156 (2005)
20. Hur, K., Santoso, S.: An improved method to estimate empirical system parallel resonant frequencies using capacitor switching transient data. IEEE Trans. on Power Delivery 21(3), 1751–1753 (2006)
21. Hur, K., Santoso, S.: Analysis and modeling of dynamic overvoltage phenomena due to transformer energizing. In: IEEE PES General Meeting, vol. 2, pp. 1126–1130 (2005)
22. Hur, K., Santoso, S.: On two fundamental signatures of shunt capacitor bank energizing for determining the relative location of switched capacitor banks. IEEE Trans. on Power Delivery 23(2), 2419–2427 (2008)
23. Chang, G., Shih, M., Chu, S., Thallam, R.: An efficient approach for tracking transients generated by utility shunt capacitor switching. IEEE Trans. on Power Delivery 21, 510–512 (2006)
24. Parsons, A., Grady, W., Powers, E., Soward, J.: A direction finder for power quality disturbances based upon disturbance power and energy. IEEE Trans. on Power Delivery 15, 1081–1085 (2000)
25. Hur, K., Santoso, S., Gu, I.: On the empirical estimation of utility distribution damping parameters using power quality waveform data. EURASIP Journal on Advances in Signal Processing 2007, Article ID:95328 (2007)

26. Hur, K., Santoso, S.: Estimation of system damping parameters using analytic wavelet transforms. IEEE Trans. on Power Delivery 24(3), 1302–1309 (2009)
27. Hauer, J.: Application of prony analysis to the determination of modal content and equivalent models for measured power system response. IEEE Trans. on Power Delivery 6, 1062–1068 (1991)
28. Scharf, L.: Statistical Signal Processing: Detection, Estimation and Time-Series Analysis. Addison Wesley, New York (1991)
29. Lobos, T., Rezmer, J., Koglin, H.: Analysis of power system transients using wavelets and Prony method. In: Proc. IEEE Porto Power Tech. Conf., Porto, Portugal, paper EMT-103 (2001)
30. Ribeiro, M., Romano, J., Duque, C.: An improved method for signal processing and compression in power quality evaluation. IEEE Trans. on Power Delivery 19(2), 464–471 (2004)
31. Santoso, S., Powers, E., Grady, W.: Power quality disturbance data compression using wavelet transform methods. IEEE Trans. Power Delivery 12, 1250–1257 (1997)
32. Dash, P., Panigrahi, B., Sahoo, D., Panda, G.: Power quality disturbance data compression, detection, and classification using integrated spline wavelet and S-transform. IEEE Trans. on Power Delivery 18(2), 595–600 (2003)
33. Ma, H., Girgis, A.: Identification and tracking of harmonic sources in a power system using a Kalman filter. IEEE Trans. on Power Delivery 11(3), 1659–1665 (1996)
34. Kim, J., Grady, W., Arapostathis, A., Soward, J., Bhatt, S.: A time domain procedure for locating switched capacitors in power distribution systems. IEEE Trans. on Power Delivery 17, 1044–1049 (2002)
35. PSCAD/EMTDC version 4.2, Manitoba HVDC Research Centre, Winnipeg, Canada
36. Bendat, J., Piersol, A.: Random data: analysis and measurement procedures. Wiley & Sons, Chichester (1986)
37. Antoulas, A.: Approximation of large-scale dynamical systems. SIAM, Philadelphia (2005)
38. Mallat, S.: A wavelet tour of signal processing, 2nd edn. Academic Press, London (1999)
39. Oppenheim, A., Schafer, R.: Discrete-time signal processing, 2nd edn. Prentice Hall, New Jersey (1999)
40. Slavič, J., Simonovski, I., Boltežar, M.: Damping identification using a continuous wavelet transform: application to real data. Journal of Sound and Vibration 262, 291–307 (2003)

Particle Swarm Optimization PSO: A New Search Tool in Power System and Electro Technology

Adel M. Sharaf and Adel A.A. El-Gammal

Abstract. Classical optimization techniques such as LP and NLP are efficient approaches that can be used to solve special cases of optimization problem in power system applications. As the complexities of the problem increase, especially with the introduction of uncertainties to the system, more complicated optimization techniques, such as stochastic programming have to be used. Particle Swarm Optimization (PSO) technique can be an alternative solution for these complex problems. Particle Swarm Optimization (PSO) is an evolutionary computational technique, (a search method based on a natural system), which was introduced by Kennedy and Eberhart in 1995. This optimization and search technique models the natural swarm behavior seen in many species of birds returning to roost, group of fish, and swarm of bees… etc. In general, there are two optimization techniques based on Particle Swarm Optimization (PSO). These two techniques are:

1. Single objective Particle Swarm Optimization SOPSO, and
2. Multi objective Particle Swarm Optimization MOPSO.

The main procedure of the SOPSO is based on deriving a single objective function for the problem. The single objective function may be combined from several objective functions using weighting factors. The objective function is optimized (either minimized or maximized) using the Particle Swarm Optimization (PSO) to obtain a single solution. On the other hand, the main objective of the Multi-Objective (MO) problem is finding the set of acceptable trade-off optimal solutions. This set of accepted solutions is called Pareto front. These acceptable solutions give more ability to the user to make an informed decision by seeing a wide range of solutions that are optimum from an "overall" standpoint. Single Objective (SO) optimization may ignore this trade-off viewpoint. This chapter has described the basic concepts of PSO and presents a review of some of the applications of PSO in power systems-based optimization problems to give the reader some insight of how PSO can serve as a solution to some of the most complicated engineering optimization problems.

1 Single Objective Particle Swarm Optimization (SOPSO)

Particle Swarm Optimization (PSO) is an evolutionary computational technique, (a search method based on a natural system), which was introduced by Kennedy and

Adel M. Sharaf and Adel A.A. El-Gammal
Centre for Engineering Studies, Energy Research, University of Trinidad and Tobago UTT, Point Lisas Campus, Esperanza Road, Brechin Castle, P.O. Box 957. Couva
E-mail: {adel.sharaf,adel.elgammal}@utt.edu.tt

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 235–294.
springerlink.com

Eberhart in 1995 [1]. This optimization and search technique models the natural swarm behavior seen in many species of birds returning to roost, group of fish, and swarm of bees… etc. The Particle Swarm Optimization (PSO) may be used to find optimal (or near optimal) solutions to numerical and qualitative problems [2] – [7]. PSOs methods are inspired by particles moving around in the defined Search-Space. The individuals in a PSO have a position and a velocity. The PSO method remembers the best position found by any particle. Additionally, each particle remembers its own previously best-found position. A particle moves through the specified solution space along a trajectory defined by its velocity, the draw to return to a previous promising search area, and an attraction towards the best location discovered by its close neighbors. Particle swarm optimization has been used for a wide range of search applications, as well as for specific optimization tasks. PSO can be easily implemented in most programming languages and has proven to be both effective and fast when applied to a diverse set of nonlinear optimization problems. PSO has been successfully applied in many areas:

- Function optimization,
- Artificial neural network training,
- Proportional and Integral Fuzzy system control, and
- Other search and optimization areas where GA can be applied.

1.1 Differences between Pso and Other Evolutionary Computation (EC) Techniques

Most of Evolutionary Techniques have the following procedure:

1. **The system is initialized with a population of random solutions and searches for optima by updating generations.**
2. **Reckoning of a fitness value for each subject. It will directly depend on the distance to the optimum.**
3. **Reproduction of the population based on fitness values.**
4. **If requirements are met, then stop.**
5. **Otherwise go back to 2.**

From the procedure, PSO shares many common points with GA.

- **Both algorithms start with a group of a randomly generated-population,**
- **Both have fitness values to evaluate the population.**
- **Both update the population and search for the optimum with random techniques.**
- **Both systems do not guarantee success.**

- Unlike other random search algorithms, each potential solution (called a particle) is also assigned a randomized velocity and then flown through the problem hyperspace.
- The most striking difference between PSO and the other evolutionary soft computing algorithms is that PSO chooses the path of cooperation over competition. The other algorithms commonly use some form of decimation, survival of the fittest. In contrast, the PSO population is stable and

individuals are not destroyed or created. Individuals are influenced by the best performance of their neighbors. Individuals eventually converge on optimal points in the problem domain.

1. The PSO traditionally does not have any genetic operators like crossover between individuals and mutation, and other individuals never substitute particles during the run. Instead the PSO refines its search by attracting the particles to positions with good solutions.
2. Particles update themselves with the internal velocity.
3. They also have memory, which is important to the algorithm.
4. Compared with genetic algorithms (GAs), the information sharing mechanism in PSO is significantly different. In GAs, chromosomes share information with each other. So the whole population moves like a one group towards an optimal area. In PSO, only $\mathbf{g}_{best}$ or $\mathbf{p}_{best}$ gives out the information to others. It is a one-way information sharing mechanism. The evolution only looks for the best solution.
5. Compared to the GA, the advantages of PSO are that PSO is easy to implement and there are few parameters to adjust.

A comparison between conventional optimization techniques and evolutionary algorithms (like genetic algorithm and PSO) is presented in Table 1 [7].

Table 1. Comparison between conventional optimization procedures and evolutionary algorithms

Property	Evolutionary	Traditional
Search space	Population of potential solutions	Trajectory by a single point
Motivation	Natural selection and Social adaptation	Mathematical properties (gradient, Hessian)
Applicability	Domain independent, Applicable to variety of problems	Applicable to a specific problem domain
Point Transition	Probabilistic	Deterministic
Prerequisites	An objective function to be optimised	Auxiliary knowledge such as gradient vectors
Initial guess	Automatically generated by the algorithm	Provided by user
Flow of control	Mostly parallel	Mostly serial
CPU time	Large	Small
Results	Global optimum more probable	Local optimum, dependant of initial guess
Advantages	Global search, parallel, speed	Convergence proof
Drawbacks	No general formal convergence proof	Locality, computational cost

1.2 The Structure of a Particle

The basic structure of any particle in a selected population consists of five components:

- $\vec{x}$, is a vector containing the current location in the solution space. The size of $\vec{x}$ is defined by the number of variables used by the problem that is being solved.
- Fitness is the quality of the solution represented by the vector $\vec{x}$, as computed by a problem-specific evaluation function.
- $\vec{V}$, is a vector containing the velocity for each dimension of $\vec{x}$. The velocity of a dimension is the step size that the corresponding $\vec{x}$ value will change into at the next iteration. Changing the $\vec{V}$ values changes the direction the particle will move through in the search space, causing the particle to make a turn. The velocity vector is used to control the range and resolution of the search.
- P_{best} is the fitness value of the best solution yet found by a particular particle, and
- $\vec{P}$ is the copy of the $\vec{x}$ for the location that generated the particle's P_{best} . Jointly, P_{best} and $\vec{x}$ comprise the particle's memory, which is used to control the particle to go back towards a definite search region.

Each particle is also aware of the current best fitness in the neighborhood for any given iteration. A neighborhood may consist of some small group of particles, in which case the neighborhoods overlap and every particle is in multiple neighborhoods. Particles in a swarm are related socially; that is, each particle is a member of one or more neighborhoods. Each individual tries to emulate the behavior of the best of its neighbors. Each individual can be thought

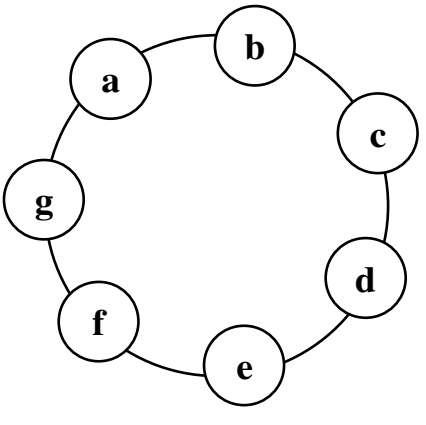


Fig. 1. Circle of Neighborhood

of as moving through the feature space with a velocity vector that is influenced by its neighbors. The particles can be viewed as being in a circle, where each particle is the center of a neighborhood made of the previuos and next particles, as shown in Fig. (1).

1.3 Basic Method

The position of each particle is represented by XY axis position; and also the velocity is expressed by V_x (the velocity of X axis) and V_y (the velocity of Y axis). Modification of the particle position is realized by the position and velocity information. Each particle knows its best value so far ($\mathbf{P_{best}}$) and its XY position. This information represents the personal experiences of each particle. Moreover, each particle knows the best value so far in the group ($\mathbf{g_{best}}$) among $\mathbf{P_{bests}}$. This information represents the knowledge of how the other particles around have performed. Namely, Each particle tries to modify its position using the following information:

- The current positions (x, y),
- The current velocities (V_x, V_y),
- The distance between the current position and $\mathbf{P_{best}}$
- The distance between the current position and $\mathbf{g_{best}}$

This modification can be represented by the concept of velocity. Velocity of each particle can be modified by the following equation:

$$\mathbf{V_{id} = W \times V_{id} + C_1 \times rand_1 \times (P_{id} - X_{id}) + C_2 \times rand_2 \times (P_{gd} - X_{id})} \quad (1)$$

Where:

- V_{id} is the value of dimension d in the velocity vector $\vec{\mathbf{V}}$ for particle i,
- C_1 is the cognitive learning selected rate,
- C_2 is the social learning selected rate,
- $rand_1$ and $rand_2$ are random values on the range [0.1],
- X_{id} is the current position of particle i along dimension d,
- W is the selected weighting factor,
- P_{id} is the location along dimension d at which the particle previously had the best fitness measure, and
- P_{gd} is the current location along dimension d of the neighborhood particle with the best fitness.

The basic concept of the PSO technique lies in accelerating each particle towards its $\mathbf{P_{best}}$ and $\mathbf{g_{best}}$ locations, with a random weighted acceleration at each step and this is illustrated in Fig. (2),

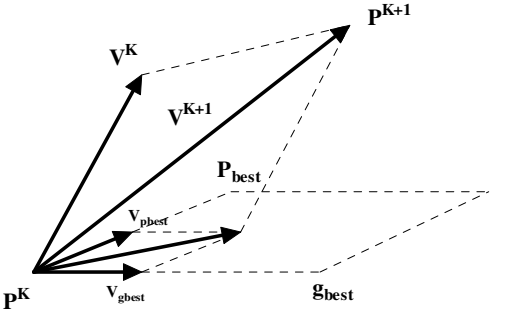


Fig. 2. Concept of modification of a searching point by PSO

Where

P_k is the current position of a particle,
P_{k+1} is its modified position,
V_K is its initial velocity,
V_{K+1} is its modified velocity,
V_{pbest} is the velocity considering its p_{best} location and
V_{gbest} is the velocity considering its g_{best} location.

Using the above concept, a certain velocity, which gradually gets close to $\mathbf{P_{best}}$ and $\mathbf{g_{best}}$, can be calculated. The current position (searching point in the solution space) can be modified by the following equation:

$$\mathbf{X}_{id} = \mathbf{X}_{id} + \mathbf{V}_{id} \tag{2}$$

1.4 The Search Algorithm

The main steps in the particle swarm optimization process are described as follows:

1. System initialized with a population of random potential solutions. Each potential solution is assigned a random 'velocity' and is called a particle. (It has position in the space; i.e., it is a point in the solution space and it has velocity). These particles are then 'flown' through the search space of potential solutions.
2. Evaluate the fitness of each particle in the swarm.
3. For every iteration, compare each particle's fitness with its previous best fitness *(P_{best})* obtained. If the current value is better than P_{best}, then set P_{best} equal to the current value and the P_{best} location equal to the current location in the d-dimensional space.
4. Compare P_{best} of particles with each other and update the swarm global best location with the greatest fitness *(g_{best})*.

5. The velocity of each particle is changed (accelerated) towards its P_{best} and g_{best}. This acceleration is weighted by a random term. A new position in the solution space is calculated for each particle by adding the new velocity value to each component of the particle's position vector.
6. Repeat steps (2) to (5) until convergence is reached based on some desired single or multiple objective criteria.

Figure (3) shows the general flow chart of the PSO algorithm.

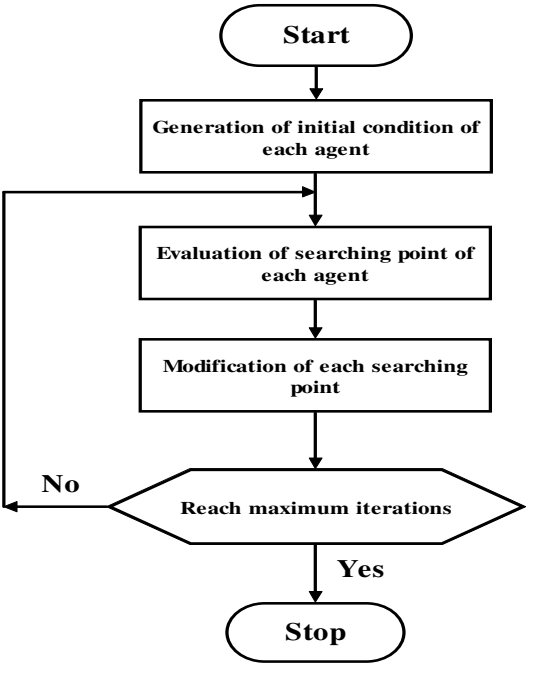


Fig. 3. The general flow chart of PSO search algorithm

There are not many parameters that need to be tuned in PSO. Here is a list of the parameters and their typical values.

- The number of particles:

  The typical range is usually 20 - 40. Actually, for most of the problems 10 particles is large enough to get good results.

- Dimension of particles:

  It is determined by the problem to be optimized; the solution space, itself, has a number of dimensions (1 or more) matching the number of variables (unknowns) in the problem. Unlike people, the PSO algorithm has no difficulty working with 4 or more dimensions.

- Range of particles:

  It is also determined by the problem to be optimized. You can specify different ranges for different dimension of particles.

- V_{max}
  It determines the maximum change one particle can take during one iteration. V_{max} is the maximum allowable velocity for the particles (i.e. in the case where the velocity of the particle exceeds V_{max}, then it is limited to V_{max}). Thus, resolution and fitness of search depends on V_{max}. If V_{max} is too high, then particles will move beyond a good solution, and if V_{max} is too low, particles will be trapped in local minima.

- Learning factors:

  The constants C_1 and C_2 in Eq. (1), termed as cognition and social components, respectively, are the acceleration constants which changes the velocity of a particle towards P_{best} and g_{best} (generally, somewhere between P_{best} and g_{best}). C_1 and C_2 are usually equal to 2. However, other settings were also used in different papers. But C_1 is usually equal to C_2 and ranges from [0- 4].

- W is called the inertia weight and controls the exploration and exploitation of the search space because it dynamically adjusts velocity. Local minima are avoided by small local neighborhoods, but faster convergence is obtained by a larger global neighborhood, and in general a global neighborhood is preferred. Synchronous updates are more costly than the asynchronous updates. Other modified error/deviation based inertia weights can be used.

- The stop condition:

  There are two possible conditions to stop the algorithm execution. These stop criteria is usually either the maximum number of iterations the PSO executes or the minimum error requirement achieved. These stop conditions depend on the problem complexity to be optimized.

2 Multi-Objective Particle Swarm Optimization (MOPSO)

In many real-life applications, multiple and often conflicting objectives need to be satisfied. Satisfying these conflicting objective functions is called Multi-Objective Optimization (MO). For example, to place more functional blocks on a chip while minimizing that chip's area and/or power dissipation are conflicting objectives that need performing a tradeoff analysis [8]. The objective of MO optimization is to find a set of acceptable solutions and present them to the user, who will then choose from them. Generally, there are two general approaches to solve Multi-Objective Optimization. The first approach lies in combining the

individual objective functions into a single composite function. Determination of a single objective is possible with methods, such as the weighted sum method, but the problem lies in the correct selection of the weights. In practice, it can be very difficult to accurately select these weights, even for someone very familiar with the problem domain. In addition, optimizing a particular solution with respect to a single objective can result in unacceptable results with respect to the other objectives [8]. The second general approach is to obtain the optimal solution. There will be a set of optimal trade-offs between the conflicting objectives, but this optimal solution is called Pareto optimal solution set or Pareto front [9]. A Pareto optimal set is a set of solutions that are non-dominated with respect to one another. While moving from one Pareto solution to another, there is always a certain amount of importance in one objective to achieve a certain amount of gain in the other. Generating the Pareto set has several advantages. The Pareto set allows the user to make an informed decision by seeing a wide range of options. The Pareto set contains the solutions that are optimum from an "overall" standpoint. SO optimization may ignore this trade-off viewpoint. This feature is useful since it provides better understanding of this system in which all the consequences of a decision with respect to all the objectives can be explored [8]. Fig (4) shows an example of the optimization of

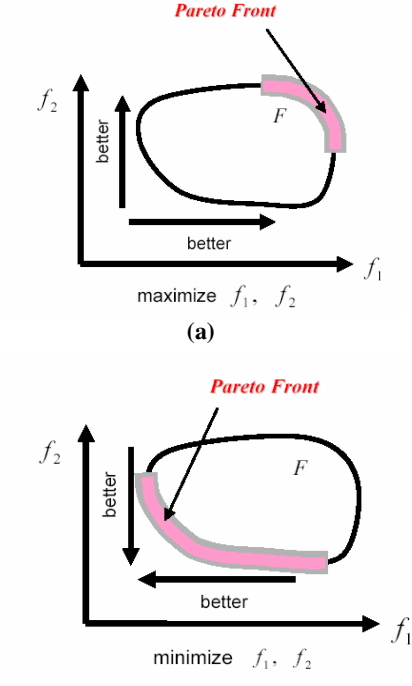


Fig. 4. Illustration of the Pareto front of a bi-objective optimization problem. Figure is extracted from [8].

two conflicting objective functions f_1 and f_2. The Pareto front that optimizes f_1 and f_2 is shown in the graph [8].

2.1 Simple Example

Consider the problem of determining the most efficient transportation mode [8]. There are two conflicting objective functions. These functions that must be optimized are:

(a) Distance covered in a day, and
(b) Energy used in the process.

The following transportation modes are assumed to be available: walking; riding a cow, a bicycle, a car, a motorcycle, a horse, an airplane, a rocket, a balloon, a boat and a scooter. These transportation modes are considered the set of available solutions. Common sense can be used to obtain all potential solutions. For example, the car will need less fuel than the motorcycle. At the same time, the car can travel longer distance. Therefore, the solution that uses the car dominates the solution that uses the motorcycle. Similarly, given the same amount of food, the horse will cover a longer distance than the cow. Therefore, the solution that uses the horse dominates the solution that uses cow. In the same manner, Fig (5) shows the set of acceptable solutions for determining the most efficient transportation mode [8].

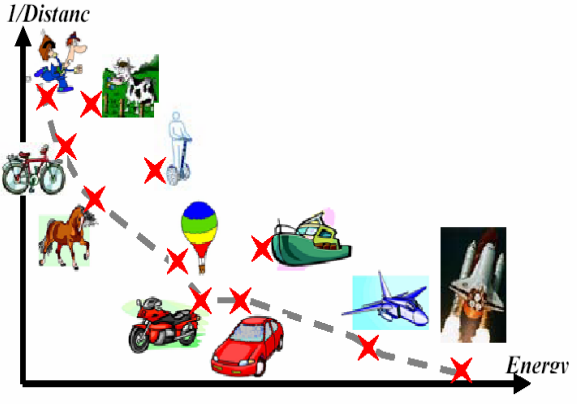


Fig. 5. Example of transportation modes, Figure is extracted from [8].

2.2 Generic Formulation of MO Optimization

The following definitions are used in the proposed Multi-Objective Optimization (MO) search algorithm:

Def. 1. The general MO problem requiring the optimization of N objectives may be formulated as follows:

Minimize

$$\vec{y} = \vec{F}(\vec{x}) = \left[\vec{f}_1(\vec{x}), \vec{f}_2(\vec{x}), \vec{f}_3(\vec{x}), ..., \vec{f}_N(\vec{x})\right]^T \quad (3)$$

$$subject \quad to \quad g_j(\vec{x}) \le 0 \qquad j = 1,2, \ldots, M \quad (4)$$

$$\text{Where: } \vec{x}^* = \left[\vec{x}_1^*, \vec{x}_2^*, \ldots, \vec{x}_P^*\right]^T \in \Omega \quad (5)$$

$\vec{y}$ is the objective vector, the $\vec{g}_i(\vec{x})$ represent the constraints and $\vec{x}^*$ is a P-dimensional vector representing the decision variables within a parameter space Ω .. The space spanned by the objective vectors is called the objective space. The subspace of the objective vectors satisfying the constraints is called the feasible space.

Def. 2. A decision vector $\vec{x}_1 \in \Omega$ is said to dominate the decision vector $\vec{x}_2 \in \Omega$ (denoted by $\vec{x}_1 \prec \vec{x}_1$), if the decision vector $\vec{x}_1$ is not worse than $\vec{x}_2$ in all objectives and strictly better than $\vec{x}_2$ in at least one objective.

Def. 3. A decision vector $\vec{x}_1 \in \Omega$ is called Pareto-optimal, if there does not exist another $\vec{x}_2 \in \Omega$ that dominates it. An objective vector is called Pareto-optimal, if the corresponding decision vector is Pareto-optimal.

Def. 4. The non-dominated set of the entire feasible search space Ω is the Pareto-optimal set. The Pareto-optimal set in the objective space is called Pareto-optimal front.

2.3 The Search Procedure for the MOPSO Algorithm

In MOPSO [8-13], a set of particles are initialized in the decision space at random. For each particle i, a position x_i in the decision space and a velocity v_i are assigned. The particles change their positions and move towards the so far best-found solutions. The non-dominated solutions from the last generations are kept in the archive. The archive is an external population, in which the so far found non-dominated solutions are kept. Moving towards the optima is done in the calculations of the velocities as follows:

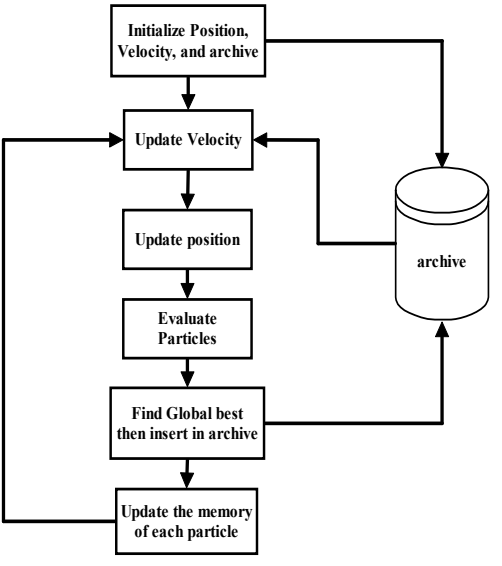
$$V_{id} = \omega \times V_{id} + C_1 \times rand_1 \times \left(P_{pd} - X_{id}\right) + C_2 \times rand_2 \times \left(P_{rd} - X_{id}\right) \quad (6)$$

$$X_{id} = X_{id} + V_{id} \quad (7)$$

Where $P_{r,d}$, $P_{p,d}$ are randomly chosen from a single global Pareto archive, ω is the inertia factor influencing the local and global abilities of the algorithm, $V_{i,d}$ is

the velocity of the particle i in the d_th dimension, c_1 and c_2 are weights affecting the cognitive and social factors, respectively. r_1 and r_2 are two uniform random functions in the range [0 , 1]. According to (7), each particle has to change its position $\mathbf{X_{i,d}}$ towards the position of the two guides $\mathbf{P_{r,d}}$, $\mathbf{P_{p,d}}$ which must be selected from the updated set of non-dominated solutions stored in the archive. The particles change their positions during generations until a termination criterion is met. Finding a relatively large set of Pareto-optimal trade-off solutions is possible by running the MOPSO for many generations.

Figure (6-a) shows the flow chart of the Multi-Objective Particle Swarm Optimization MOPSO. Also fig (6-b) explains the procedure of the Multi-Objective Particle Swarm Optimization MOPSO using pseudo code.



(a)

```
1.  MOPSO
2.  {
3.      Init _Pop();
4.      Init _Velocity();
5.      Evaluate_ Pop();
6.      Update_ Fbest();
7.      Update_ Pbest();
```

Fig. 6. Procedure of the MOPSO (a) Flow chart of the algorithm, (b) Pseudo code of MOPSO

```
8.              Insert_ nodom();
9.              Gbestpos = rnd(0,nodomfileSize)
10.             for (i=1 to MAXCYCLES)
11.             {
12.                 for (j=0 to MAXPARTICLES)
13.                 {
14.                         Update_Velocity();
15.                         Update_Particle();
16.                 }
17.                 Keeping();
18.                 Evaluate_ Pop();
19.                 Update_ Fbest();
20.                 Update _Pbest();
21.                 Insert_ nodom();
22.                 Gbestpos = rnd(0,nodomfileSize)
23.             }
24.             Print Statistics();
25.             Generate Outfile();
26.     }
```

(b)

Fig. 6. (*Continued*)

Init Pop();
The particle swarm is initialized with random values corresponding to the ranges of the decision variables. These values are dependent on the test functions.

Init Velocity();
The velocities are initialized with zero values.

Check the feasibility of each particle. If the particle does not satisfy the constraints, then regenerate it.

Evaluate Pop();
The swarm is evaluated using the corresponding objective functions.

Update Fbest();
The fitness vectors are updated (Evaluate the multi-objective fitness value of each particle and save it in vector form).

As we are dealing with multi-objective optimization, these vectors store the values of each decision variable, whereby the particles obtain the best values in a Pareto sense.

At this stage of the algorithm, these vectors are filled with the results of the initial particle evaluations.

Update Pbest();
Analogously, these values are copied in the pbest vectors.

Insert nodom();
Calculate the multi-objective fitness values of each particle and check its Pareto optimality. Store the non-dominated particles in the Pareto archive. If the specific constraint does not exist for the archive, the size of the archive will be unlimited (all non-dominated particles are inserted in the grid; i.e., in the external file).

Gbestpos = rnd (0,nodomfileSize)
The global gbest particle is randomly selected.

//The flight cycle starts here //

Update Velocity();
The velocity of each particle is updated, using Eq. (4.4).

Two Pareto solutions are chosen randomly for $\mathbf{P_{r,d}}$, $\mathbf{P_{i,d}}$ from the Pareto archive.

Update Particle();
The position of each particle is also updated using Eq. (4.4)

Keeping();
The keeping operation is carried out to maintain the particles into the allowable range values. Then, the particles are mutated.

If the particle does not remain within the feasible solution region, it is discarded and mutated again.

Evaluate Pop(); Update Fbest(); Update Pbest();
The particles are evaluated, the fitness, and pbest vector are, if appropriate, updated.

Insert nodom();
As the particles move in the search space because they have changed positions, the dominance of each particle is verified and, if appropriate, they are inserted in the grid.

(a) Check the Pareto optimality of each particle. If the fitness value of the particle is non-dominated when compared to the Pareto optimal set in the archive, save it into the Pareto archive.
(b) In the Pareto archive, if a particle is dominated by a new one, then discard it.

Gbestpos = rnd(0,nodomfileSize)
Then, the new gbest is randomly selected. Two Pareto solutions are chosen randomly for $p_{p;d}$ and $p_{r;d}$ from the Pareto archive.

Repeat the cycle until the number of generations reaches a given n.

// the end of the cycles //

Print Statistics(); Generate Outfile();
Print the statistics and generate an output file, which contains the non-dominated particles.

2.4 Test Functions

In order to validate the proposed search algorithm of MOPSO, the following well-known test functions are used:

Test problem 1:

$$\begin{aligned} &\text{Minimize} \quad f_1(x_1, x_2) = x_1 \\ &\text{Minimize} \quad f_2(x_1, x_2) = \frac{(1+x_2)}{x_1} \\ &\text{subject to} \quad 0.1 \le x_1 \le 1,\ 0 \le x_2 \le 5 \end{aligned} \tag{8}$$

Test problem 2:

$$\begin{aligned} &\textbf{Maximize} \quad f_1(x_1, x_2) = 1.1 - x_1 \\ &\textbf{Maximize} \quad f_2(x_1, x_2) = 60 - \frac{(1+x_2)}{x_1} \\ &\textbf{subject to} \quad 0.1 \le x_1 \le 1,\ 0 \le x_2 \le 5 \end{aligned} \tag{9}$$

Test problem 3:

$$\begin{aligned} &\textbf{Minimize } f_1(x_1, x_2) = x_1^2 + x_2^2 \\ &\textbf{Minimize } f_2(x_1, x_2) = (x_1 - 5)^2 + (x_2 - 5)^2 \\ &\textbf{subjectto} -5 \le x_1 \le 10, -5 \le x_2 \le 10 \end{aligned} \tag{10}$$

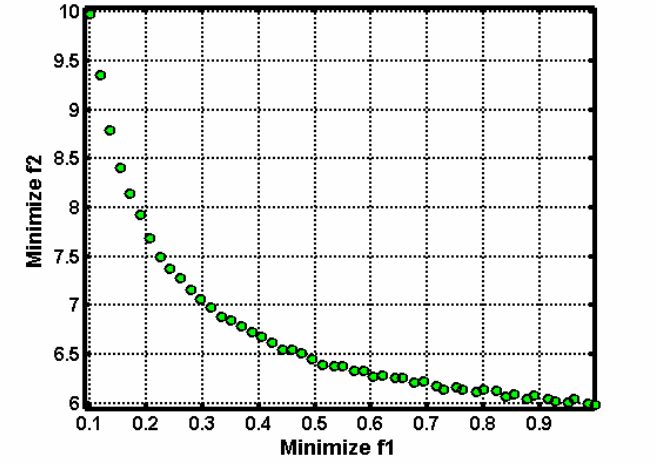


Fig. 7. Objective space of test problem 1

Test problem 4:

$$\begin{aligned} &\text{Minimize} \quad f_1(x_1,x_2) = \left(x_1^2 + x_2^2\right)^{\frac{1}{8}} \\ &\text{Minimize} \quad f_2(x_1,x_2) = \left((x_1 - 0.5)^2 + (x_2 - 0.5)^2\right)^{\frac{1}{4}} \\ &\text{subject to} \quad -5 \le x_1 \le 10, -5 \le x_2 \le 10 \end{aligned} \tag{11}$$

Figures (7)-(10) show the Pareto fronts (or objective space) and the Pareto sets (input space) for each test function, respectively.

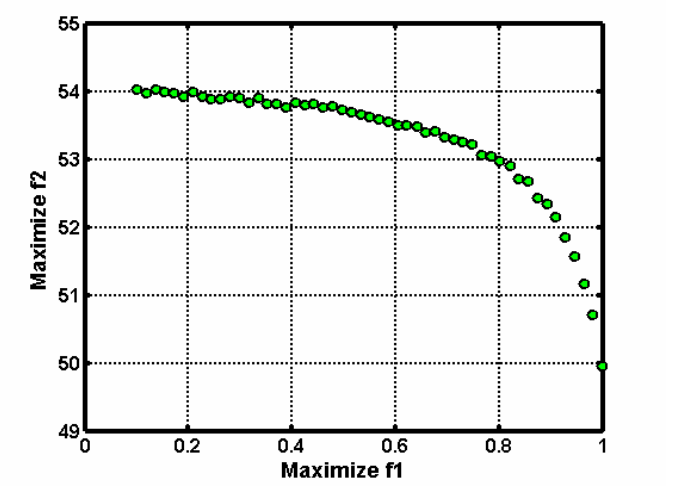


Fig. 8. Objective space of test problem 2

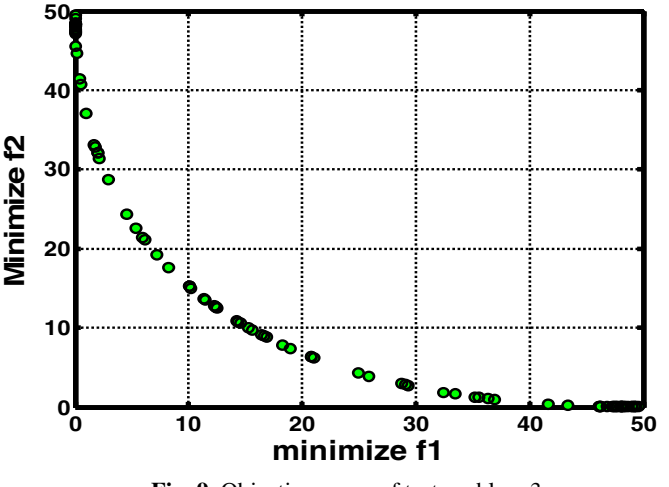


Fig. 9. Objective space of test problem 3

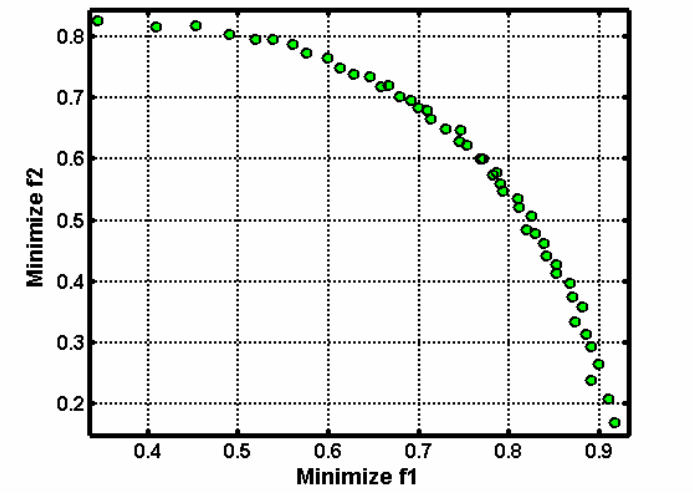


Fig. 10. Objective space of test problem

3 Discrete Particle Swarm Optimization (DPSO) [14-15]

The particle swarm search algorithm works by adjusting trajectories through manipulation of each coordinate of a particle in the d-dimensional space. However, many optimization problems are set in a space featuring discrete, discontinuous qualitative distinctions between variables and between levels of variables. In the binary version of the PSO, the trajectories are changes in the probability that a coordinate will take on binary value (0 or 1). Therefore, the main difference between the original PSO and the DPSO is equation (3) replacing equation (2).

$$\begin{aligned} &if\ (rand() < S(v_{id}))\ then\ \ x_{id} = 1; \\ &\qquad\qquad Else\ \ x_{id} = 0. \end{aligned} \tag{12}$$

Where S (v) is a sigmoid limiting transformation function:

$$S(v) = \frac{1}{1+e^{-v}} \tag{13}$$

and rand( ) is a quasi-random number selected from a uniform distribution in [0.0, 1.0]. In the discrete version, V^{max} is retained, that is $|v_{id}| < V^{\max}$ which simply limits the ultimate probability that bit x_{id} will take on a binary value. A smaller V^{max} will allow a higher mutation rate.

4 PSO Applications in Power System

Solving an optimization problem is one of the common scenarios that occur in most engineering applications. Classical techniques such as LP and NLP are efficient approaches that can be used in special well defined cases. As the complexities of the problem increase, especially with the introduction of severe uncertainties to the system, more complicated optimization techniques, such as stochastic programming or DP have to be used. However, these analytical methods are not easy to implement for most of the real-world problems [16]. In fact, for many problems, the curse of dimensionality makes the approach unfeasible to implement. The above issues are of particular importance when solving optimization problems in a power system. As a highly nonlinear, non-stationary system with noise and uncertainties, a power network can have a large number of states and parameters. Implementing any of the classical analytical optimization approaches might not be feasible in most of the cases. On the other hand, PSO can be an alternative and effective solution. It is a stochastic-based search technique that has its roots in artificial life and social psychology, as well as in engineering and computer science. This section will present some of the applications of PSO in power systems-based optimization problems to give some insight of how PSO can serve as a solution to some of the most complicated engineering optimization problems [17-47].

4.1 Efficient Operation of Induction Motor Drives [17]

This application presents an optimization technique for an efficient controller for the three-phase induction motors. Multi Objective Particle Swarm Optimization MOPSO technique is implemented to tackle all the conflicting goals that define the search for the optimality problem. The PSO search deals with two main conflicting objective functions. These conflicting functions are: Maximizing the operating efficiency of the drive system for a given mechanical load, and maximizing the equivalent power factor of the induction motor for start up and steady state operation. In addition, the optimization ensures that maximum allowable stator current constraints are not exceeded. The proposed technique are based on the principle that the flux level in a machine can be adjusted to give the required trade-off solution of maximum efficiency and maximum power factor for a given value of speed and load torque. The optimum flux levels are function of the machine load and speed requirements. Simulation results show that considerable efficiency and power factor improvements are achieved using MOPSO when compared with the Field Oriented Control (FOC) and Constant Voltage to Frequency Ratio based Control (CVFRC).

4.1.1 The Objective Functions

The following definitions are useful in subsequent analyses, the per-unit frequency is

$$a = \frac{\omega_e}{\omega_b} = \frac{\omega_s + \omega_r}{\omega_b} \tag{14}$$

The slip is defined by:

$$s = \frac{\omega_s}{\omega_e} = \frac{\omega_s}{\omega_s + \omega_r} \tag{15}$$

The rotor current is given by:

$$I'_r = \frac{\phi_m}{\sqrt{\left(\frac{r'_r}{sa}\right)^2 + X'^2_{lr}}} \tag{16}$$

The electromagnetic torque is given by:

$$T_e = \frac{\left(\frac{r'_r}{sa}\right)}{\left(\frac{r'_r}{sa}\right)^2 + X'^2_{lr}} \phi^2_m \tag{17}$$

The stator current is related to the air gap flux and the electromagnetic torque as:

$$I_s = \sqrt{\left(s_1\phi_m + s_2\phi_m^3 + s_3\phi_m^5\right)^2 + C_L \frac{T_e^2}{\phi_m^2}} \tag{18}$$

Where

$$C_L = 1 + 2 \times \frac{x'_{lr}}{x_m} \tag{19}$$

The air gap flux is related to the electromagnetic torque as:

$$\phi_m = \sqrt{\frac{sa}{r'_r}} \sqrt{\left(\frac{r'_r}{sa}\right)^2 + x'^2_{lr}} \sqrt{T_e} \tag{20}$$

r_s :stator resistance, r_r :rotor resistance, X_{ls} :stator leakage reactance, X_{lr} :rotor leakage reactance, a, ω_e :Supply frequency, S :slip, ω_r :rotor speed, ω_b :base speed, ω_s: slip speed, ϕ_m :air gap flux, I_s: stator current, I_r :rotor current, T_e :electromagnetic torque, T_L:load torque, k_e, k_h :eddy current and hysteresis coefficients, c_{str} :stray losses coefficient, $c_{f\omega}$:mechanical losses coefficient, s_1,s_2, s_3 :magnetizing curve coefficients. The power efficiency is defined as the output power divided by the electric power supplied to the stator (inverter losses are included). After some mathematical manipulations, the efficiency is expressed as:

$$\eta = \frac{(T_e \times \omega_r)}{\left((r_s I_s^2 + r_r' I_r'^2) + \left(k_e(1+s^2)a^2\ \phi_m^2\right) + \left(k_h(1+s)\,a\ \phi_m^2\right) + \left(C_{str}\,\omega_r^2\ I_r'^2\right) + + \left(C_{f\omega}\ \omega_r^2\right) + \left(K_{1inv}\, i_s^2 + K_{2inv}\, i_s\right) + \left(T_l \times \omega_r\right)\right)} \tag{21}$$

Also the conventional power factor is defined as the input power divided by the apparent power. After some algebraic manipulations, the power factor can be expressed as:

$$PF = \frac{(K_1 \times K_2 + K_3)}{\sqrt{(K_2^2 + K_3^2)(K_1^2 + 1)}} \tag{22}$$

Where

$$K_1 = \frac{\omega_s \times X_{rr}}{\omega_b \times r_r} \tag{23}$$

$$K_2 = r_s \times K_1 + \frac{\omega_e}{\omega_b}\left(X'' + \frac{X_m^2}{X_{rr}}\right) \tag{24}$$

$$K_3 = r_s - \frac{\omega_e}{\omega_b}(X'' \times K_1) \tag{25}$$

Where X'' is the sub transient reactance:

$$X'' = X_{ls} + \frac{X_m X_{lr}}{X_m + X_{lr}} \tag{26}$$

For optimal operation, three objective functions that can affect the operation optimization of three phase induction motors had been chosen. These two objective functions are:

1. Efficiency (to maximize);
2. Power factor (to maximize);

In addition, the optimization ensures that maximum allowable stator current constraints are not exceeded. Figure (11) is a plot of input power versus stator current. As illustrated in this figure, the stator current and the input power are minimized almost simultaneously. Therefore, in practice the Maximum Efficiency ME and the Minimum Stator Current MSC are not conflicting objective functions. So, when the efficiency is maximized, the stator current is minimized and the torque per ampere is maximized. So, the maximum efficiency and power factor will be considered as the main conflicting objective functions. However, a good operation should represent the right compromise among different objectives but the problem consists in searching this "compromise." The only tool to solve this problem is represented by the multi objective approach. This approach allows us to investigate how each single-objective and multi objective problem affects the results in terms of performance and independent variables and, above all, allows us to have a wide range of alternative solutions among which the operator can choose a better solution.

4.1.2 Digital Simulation Results

The MOPSO algorithm has been applied to optimize the operation of a three-phase, 380 V, 1-HP, 50 Hz, 4-pole, squirrel cage induction motor. The motor

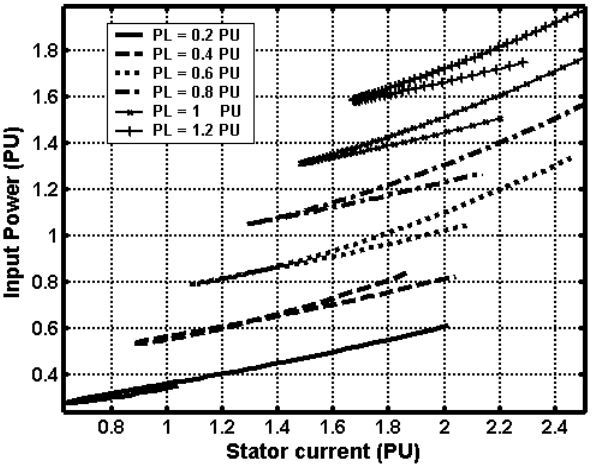


Fig. 11. Input power versus stator current at rated speed and various load powers.

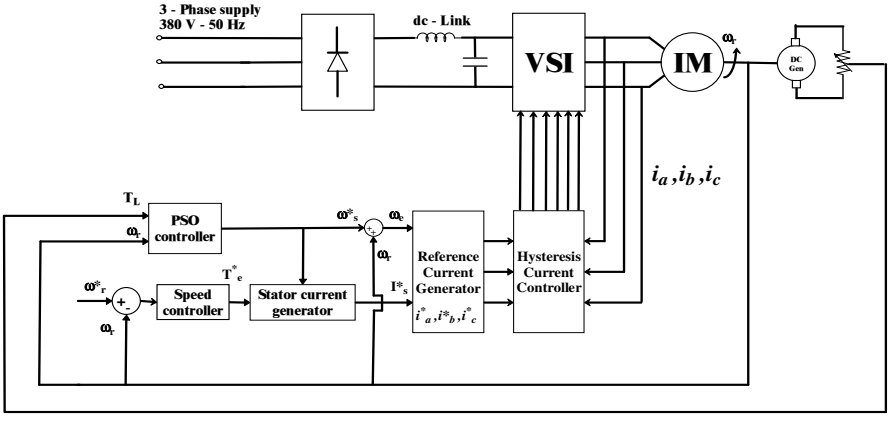
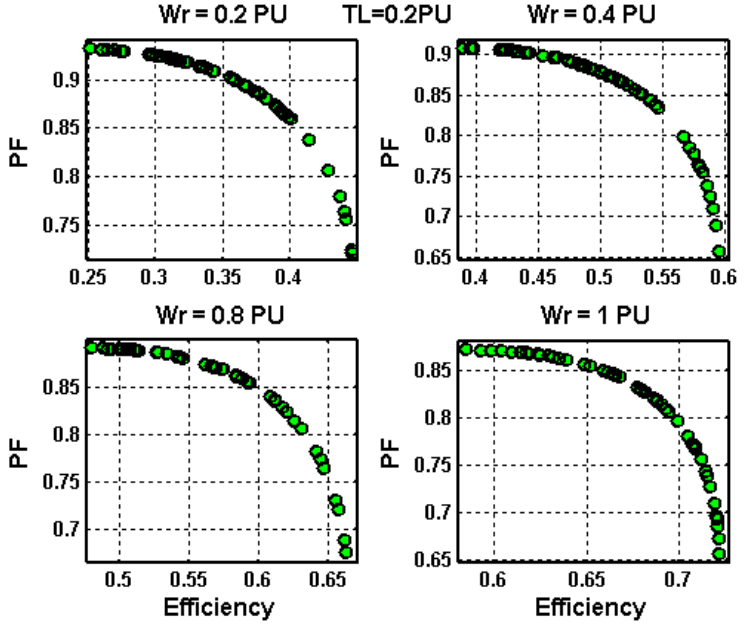


Fig. 12. The proposed drive system using PSO based on the loss model controller

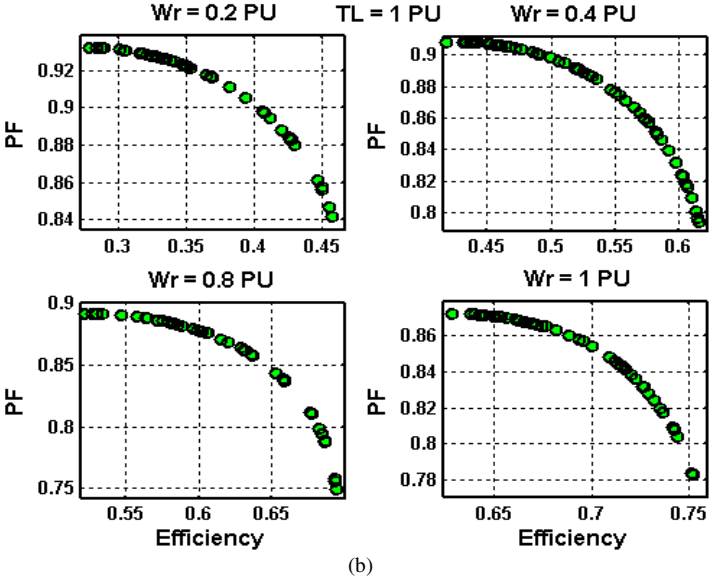
parameters in per units are R_s=0.0598, X_{ls}=0.0364, X_m=0.8564, X_{lr}=0.0546, R_r=0.0403, K_e=0.0380, K_h=0.0380, C_{str}=0.0150, C_{fw}=0.0093, S_1=1.07, S_2=-0.69, S_3=0.77.

The block diagram of the optimization process based on PSO is shown in fig. 12. In the proposed controller, the PSO algorithm receives the rotor speed and load torque, and then the PSO controller determines the slip frequency at which the optimal fitness function occurs at that rotor speed and load torque. As stated before, this part of simulation will consider the Maximum Efficiency and the Maximum Power Factor (ME and MPF) problem to be solved using Multi-Objective Particle Swarm Optimization. Figure 13 shows the solution to this problem and shows the

Pareto fronts of the problem for different levels of rotor speed ω_r = 0.2, 0.4, 0.8, 1 PU and different levels of load torque T_L = 0.2, and 1 PU.



(a)



(b)

Fig. 13. Pareto front of ME and MPF problem for different levels of rotor speed ω_r = 0.2, 0.4, 0.8, 1 PU and different levels of load torque: T_L=0.2 PU (b) T_L = 1 PU

Table 2 shows the solution limits of the efficiency and the power factor for each operating point. This range of solutions which is called the Pareto front enables the operator to choose the best compromise solution.

Table 2. The limits pf the Pareto front of the two conflicting objective functions

(a)

ω_r = 0.2 PU				
T_L (PU)	Efficiency		Power Factor	
	Minimum	Maximum	Maximum	Minimum
0.2	0.2521	0.4463	0.9323	0.6806
0.4	0.2647	0.4768	0.9323	0.7184
0.6	0.2742	0.4792	0.9323	0.7705
0.8	0.2702	0.4725	0.9323	0.7741
1	0.278	0.4636	0.9323	0.8165

(b)

ω_r = 0.4 PU				
T_L (PU)	Efficiency		Power Factor	
	Minimum	Maximum	Maximum	Minimum
0.2	0.3894	0.5957	0.9082	0.6328
0.4	0.4021	.6257	0.9082	0.6647
0.6	0.4154	0.6309	0.9082	0.7033
0.8	0.4335	0.6271	0.9079	0.7271
1	0.4182	0.6195	0.9082	0.7532

(c)

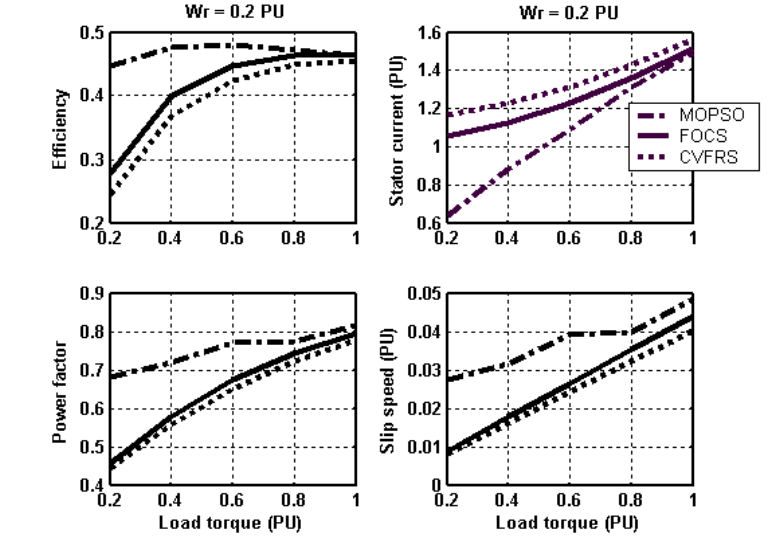
ω_r = 0.6 PU				
T_L (PU)	Efficiency		Power Factor	
	Minimum	Maximum	Maximum	Minimum
0.2	0.4804	0.6649	0.892	0.6342
0.4	0.4984	0.6937	0.892	0.6375
0.6	0.5132	0.7005	0.892	0.6701
0.8	0.5183	0.699	0.892	0.7224
1	0.523	0.6943	0.892	0.736

(d)

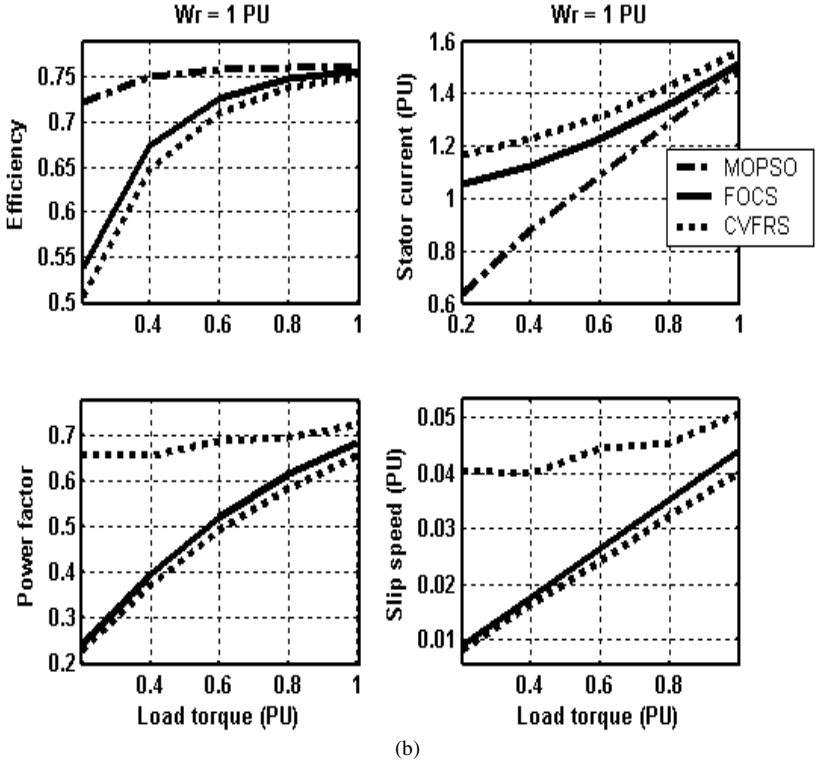
ω_r = 0.8 PU				
T_L (PU)	Efficiency		Power Factor	
	Minimum	Maximum	Maximum	Minimum
0.2	0.5475	0.7013	0.8806	0.6221
0.4	0.5681	0.7295	0.8806	0.6676
0.6	0.5808	0.7377	0.8805	0.6542
0.8	0.5793	0.7385	0.8806	0.7124
1	0.5776	0.7353	0.8805	0.7165

(e)

ω_r = 1 PU				
T_L (PU)	Efficiency		Power Factor	
	Minimum	Maximum	Maximum	Minimum
0.2	0.5846	0.7217	0.8722	0.6563
0.4	0.6093	0.7506	0.8722	0.6529
0.6	0.6221	0.7597	0.8722	0.6853
0.8	0.6277	0.7617	0.8722	0.6928
1	0.6283	0.7603	0.8722	0.7257



(a)



(b)

Fig. 14. The comparison between CVFRS, FOCS, and MOPSO for different levels of load torque T_L and different levels of rotor speed: (a) ω_r =0.2 PU (b) ω_r = 1 PU

Figure 14 shows the comparison between the constant voltage to frequency ratio strategy CVFRS, the field oriented control strategy FOCS, and the available solutions from MOPSO at different levels of load torque and rotor speed $\omega_r = 0.2$, and 1 PU respectively. The solution that has maximum efficiency is selected from the Pareto front to achieve this comparison. Keep in mind that, the operator can choose the compromised solution from the Pareto front depending on the desired efficiency and power factor. It is obvious from fig. 14 that the stator current is minimized using MOPSO. In addition, there is a great improvement in efficiency and power factor using MOPSO when compared with other strategies especially at light loads. It is noted that there is a very poor PF obtained using conventional methods (field oriented control strategy and constant voltage to frequency ratio) especially at light loads. It is obvious from fig 4 that the efficiency improvement has a noticeable value especially at light loads and rotor speed $\omega_r = 0.2$ PU that can be as high as 80 % using CVFRS. This difference decreases to 65% at rotor speed $\omega_r = 1$ PU. On the other hand, the power factor improvement reaches 55 % at light load and rotor speed $\omega_r = 0.2$ PU. Whereas this improvement reaches to 160 % at rotor speed $\omega_r = 1$ PU and light load. The performance difference between the Field Oriented Control (FOC) and the proposed control strategies MOPSO comes from the chosen value of the slip frequency and the air gap flux for each strategy. For example, at loading condition of rotor speed equal to 0.2 PU and load torques varying from 0.2 PU to 1 PU, the values of slip frequency based on FOCS vary from 0.01 PU to 0.042 PU. On the other hand, the values of slip frequency based on MOPSO vary from 0.0301PU to 0.0524 PU. This difference in slip frequency and air gap flux causes the difference in performance of each control strategy. The difference in the flux level of the (FOCS) controller comes from the choice of the level of the magnetization current command. The command of magnetization current is set to a constant value, which produces rated torque at rated stator flux. At light loads, the constant chosen value of magnetization current fails to choose the optimal flux level. On the other hand, this value yields the optimal flux level at rated loads. The proposed strategies overcome this problem and successfully choosing the optimal flux level especially at light loads.

4.2 PSO Self-regulating Modulated Power Filter Compensator Scheme for Electric Distribution Networks [18]

This application presents a novel Modulated Power Filter and Compensator (MPFC) scheme for combined voltage stability, energy conservation, loss reduction, power factor correction, and power quality enhancement of electric distribution grid systems based on Multi-Objective Particle Swarm Optimization (MOPSO). The MPFC scheme was developed by the First Author to vary the shunt power filter equivalent admittance, and modify the reactive power flow to the distribution network. The filter dynamic switching is achieved using two complementary switching pulses generated by a Sinusoidal Pulse Width Modulation (SPWM) control strategy and regulated by a tri-loop dynamic error

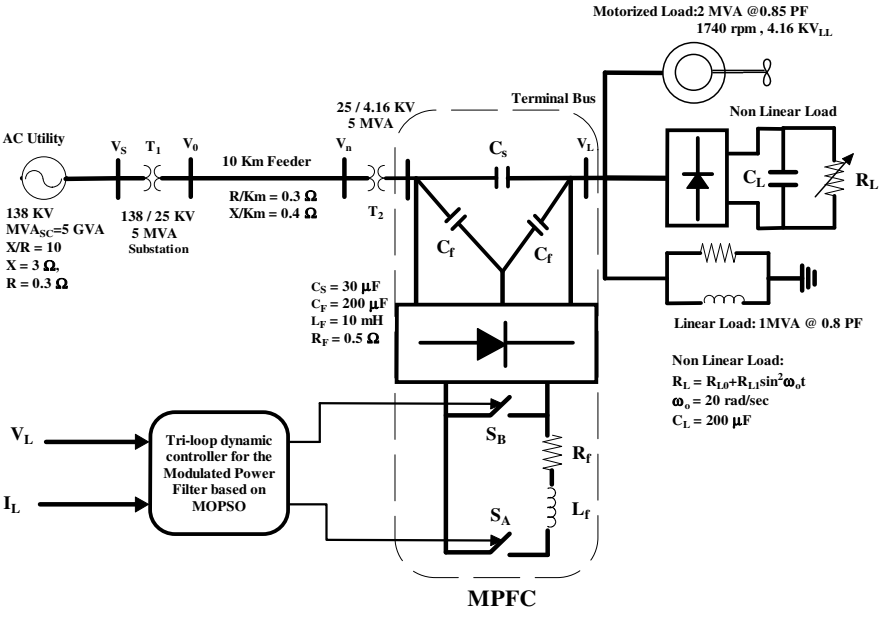


Fig. 15. Three-phase sample study ac system with the proposed Modulated Power Filter Compensator

driven controller comprising three time decoupled control loops, a minimum RMS source current dynamic loop, voltage stabilization loop, and synthesized dynamic power loop. The MOPSO technique is used to find the optimal control settings that control the input control modulation signal to the SPWM activation/ triggering block that minimizes the distribution feeder current, hence reducing feeder losses, bus voltage deviations, and ensuring distribution feeder capacity release.

4.2.1 System Description

The sample study system comprises three-phase ac utilization grid system, short feeder, hybrid electric load including a motorized load (3-phase induction motor), non-linear load, and linear load as shown in fig (15). The tri-loop error-driven dynamic controller is a novel structure developed by the First Author and used to modulate the power filter compensator PWM switching. The global error is the summation of the three loop individual errors including voltage stability, current limiting and synthesize dynamic power loops. The global error signal is input to the self tuned variable structure sliding mode controller. The (per-unit) three dimensional-error vector (e_v ,e_I, e_p) is governed by the following equations:

$$e_v(k) = \frac{V_{Lref}(k) - V_L(k)\left(\frac{1}{1+ST_0}\right)}{V_{base}} \tag{27}$$

$$e_I(k) = \frac{I_{Lref}(k) - I_L(k)\left(\frac{1}{1+ST_0}\right)}{I_{base}} \tag{28}$$

$$e_p(k) = \left(\frac{I_L(k)}{I_{base}} \times \frac{V_L(k)}{V_{base}}\right)\left(\frac{1}{1+ST_0}\right) - \left(\frac{I_L(k)}{I_{base}} \times \frac{V_L(k)}{V_{base}}\right)\left(\frac{1}{1+ST_0}\right)\left(\frac{1}{1+ST_0}\right) \tag{29}$$

The total error e_t (k) at a time instant:

$$e_t(k) = \gamma_v e_v(k) + \gamma_I e_I(k) + \gamma_p e_p(k) \tag{30}$$

The solid-state switches (S_1, S_2) are usually (GTO, IGBT/bridge, MOSFET/bridge, SSR, TRIAC) turns "ON" when a gating pulse g(t) is applied by the activation switching circuit as shown in fig. (15). Removing the pulse will turn the solid-state switch "OFF"

$$T_{S/W} = \frac{1}{f_{s/w}} = t_{on} + t_{off} \tag{31}$$

Where: $f_{s/w}$ is switching frequency, and $0<t_{on}<T_{S/W}$.

The novel filter and compensator scheme is a low-cost attractive solution for both distribution and utilization radial circuits, feeding a nonlinear load. Figure (16) depicts the self tuned variable structure sliding mode controller developed by the First Author for adjusting the switching duty-cycle-ratio (α) based on Multi Objective Particle swam Optimization searching technique MOPSO. The effective reactance of the combined hybrid fixed capacitors and the modulated tuned arm filter depends on the duty cycle and the frequency of the SPWM output which in turn is a function of the self tuned variable structure sliding mode controller output. The output of the SPWM generator is a train of pulses with variable duty cycles and constant frequency. The degree of reactive compensation is dependent

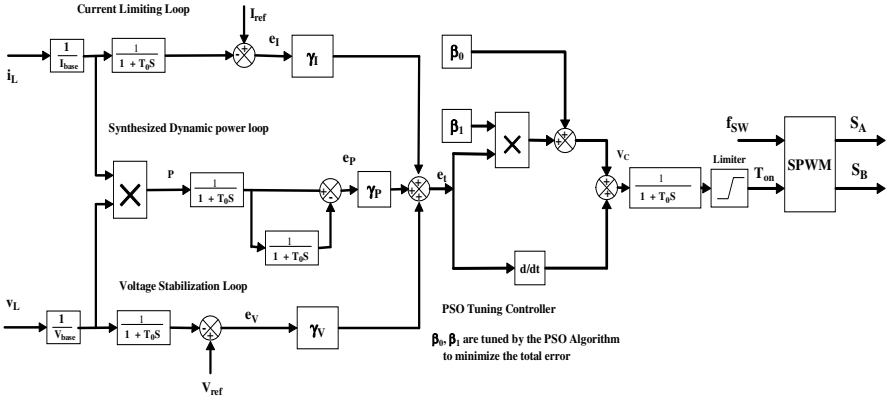


Fig. 16. Tri-loop self tuned variable structure sliding mode dynamic controller for the Modulated Power Filter MPFC

on the duty cycle of the generated pulses. This would in turn vary the effective reactance of the hybrid power filter.

The system control voltage has the following form in the time domain:

$$V_c(t) = \beta_o + \beta_1 e_t(t) + \frac{d}{dt}\left(e_t(t)\right) \tag{32}$$

The MOPSO searching algorithm is implemented for tuning the gains (β_0, β_1) to minimize the system objective functions. The selected objectives functions in this paper are to minimize a stated number of objective functions using PSO algorithm are defined by the following:

1. Minimize the absolute voltage deviations:

$$\begin{aligned} \Delta V_o &= V_S - V_o \\ \Delta V_n &= V_S - V_n \\ \Delta V_L &= V_S - V_L \end{aligned} \tag{33}$$

2. Minimize the Distribution Feeder total active power Losses:

$$J_{Ploss} = I_L^2\left[R_{T1} + R_f + R_{T2}\right] \tag{34}$$

3. Minimize the Distribution Feeder total reactive power losses:

$$J_{Qloss} = I_L^2\left[X_{T1} + X_f + X_{T2}\right] \tag{35}$$

4. Minimize the absolute total error deviations:

$$|e_t| = \left|\gamma_I e_I + \gamma_v e_v + \gamma_p e_p\right| \tag{36}$$

4.2.2 Digital Simulation Results

Matlab-Simulink Software environment was used to design, test, and validate the effectiveness of the proposed novel MPFC device and the associated dynamic SPWM controller based on SOPSO and MOPSO search optimization techniques. Table (3) shows the main objective functions versus the Tuned controller Gains based SOPSO and MOPSO control schemes, SOPSO obtains a single global or near optimal solution based on a single weighted objective function. The weighted single objective function combines several objective functions using specified or selected weighting factors as follows:

$$weighted\ objective\ function = \alpha_1 J_1 + \alpha_2 J_2 + \alpha_3 J_3 + \alpha_4 J_4 \tag{37}$$

Where $\alpha_1 = 0.25$, $\alpha_2 = 0.25$, $\alpha_3 = 0.25$, $\alpha_4 = 0.25$, are selected weighting factors.

J_1 : Minimize the voltage deviations,
J_2 : Minimize the Distribution Feeder total active power Losses,
J_3 : Minimize the Distribution Feeder total reactive power Losses,
J_4 : Minimize the absolute total error deviations.

Table 3. The main objective functions versus the Tuned controller Gains based SOPSO and MOPSO control scheme

	β_0	β_1	J_1 Minimize the voltage deviations (PU)	J_2 Minimize the Distribution Feeder total active power Losses (PU)	J_3 Minimize the Distribution Feeder total reactive power losses (PU)	J_4 Minimize the absolute total error deviations
SOPSO	0.7498	19.5937	0.0462	0.0612	0.07478	0.21746
MOPSO	0.5108	42.3264	0.019	0.0771	0.0273	0.1273
	0.1167	26.3051	0.0422	0.0554	0.07080	0.1024
	0.8393	10.2121	0.0233	0.0396	0.0371	0.2788
	0.5002	33.6397	0.0517	0.0662	0.0352	0.1398
	0.6539	41.9221	0.0514	0.0537	0.0976	0.1597
	0.8127	21.0800	0.0360	0.0456	0.0837	0.2323

Table 4. System behavior comparison with and without MPFC based SOPSO and MOPSO optimization technique

	without the MPFC	with the MPFC-with SOPSO Optimization Technique	with the MPFC-with MOPSO Optimization Technique
RMS Voltage (PU)	0.863454	0.93567	0.9634542
Power Factor	0.456798	0.9278645	0.959835
Maximum Transient Voltage – Over/Under Shoot (PU)	0.129756	0.093454	0.092563
Maximum Transient Current – Over/Under Shoot (PU)	0.09826	0.083654	0.073597
RMS Current (PU)	0.674564	0.4787465	0.429875
Active Power Losses (PU)	0.113746	0.0687231	0.0498576
Reactive Power Losses (PU)	0.148575	0.072543	0.598324

On the other hand, the MOPSO finds the set of acceptable (trade-off) near Optimal Solutions. This set of accepted solutions is called Pareto front, as shown in Table (3), MOPSO obtains six near optimal point as a Pareto front. These acceptable trade-off multi level solutions give more ability to the user to make an informed decision by seeing a wide range of near optimal selected solutions. Table (4) shows system behavior comparison with and without MPFC based SOPSO and MOPSO optimization technique, Comparing the dynamic response results of the two study cases, with and without the hybrid modulated power filter compensator, it is quite apparent that the hybrid modulated power filter compensator highly improved the ac system dynamic performance from a general power quality point of view. The effect of the MPFC is noticeable where it highly improved the power factor by reducing the amount of reactive power drawn from the supply, maintaining the system dynamic stability under sudden disturbances and recovering

faster from any inrush transient state. The MPFC had a great impact on the supply power factor improving it from 0.25 to around 0.92 which is highly desired. Moreover, the load power factor also improved from 0.2 to a value of .85. The simulation results show that the power factor can be effectively improved (from 0.35 to 0.9) when the novel dynamic compensator is used in the unbalanced fault case and from 0.4 to 0.85 in the balanced fault case. In addition, the phase voltage can maintained around 1pu and the transient over-voltages and surge type inrush currents are also damped. By comparing the results of all sample simulation cases, it is concluded that the novel dynamic MPF-Facts device with the SPWM dynamic controller developed by the first author is an attractive low cost and efficient Voltage stabilization and power factor correction device that also improve power quality and efficient-utilization of three phase –four wire Residential and Commercial Loads.

4.3 Optimal Design of Hybrid Power Filter Compensator [19]

This application presents a novel algorithm for a discrete search optimization and an approach to solve the problem of the hybrid power filter compensator with the design of C-type filter and fixed capacitor bank using Discrete Multi Objective Particle Swarm Optimization MOPSO method. This novel optimization approach, a Multi Objective Particle Swarm Optimization MOPSO method is implemented to tackle a number of conflicting search goals that define the complex optimal filter design problem. The paper presents the selection with conflicting objective functions and a compromising selection criterion:

1. Minimum change in the fundamental frequency load bus voltage under steady state conditions,
2. Minimum feeder current for maximum AC system grid capacity release,
3. Minimum fundamental frequency utilization feeder active and reactive power losses, due to reduced fundamental RMS current magnitude,
4. Minimum dominant harmonic current penetration into the host electric grid system,
5. Maximum harmonic current absorption by the hybrid harmonic power filter with the fixed capacitor bank,
6. Minimum harmonic voltage distortion at the point of common coupling or load bus.

4.3.1 Hybrid Shunt Power Filter Compensator Design

Figure 17 shows the single line diagram of a sample electric AC grid system. This electric AC grid utility-power filter installation is using the fixed C-type power filters structure shown in Fig. 19. The equivalent circuit of the system for the fundamental frequency component is shown in fig. 20. This equivalent circuit is used to predict the values of some objective functions where: $S = j\omega_1$ and ω_1=377 rad/sec. is the fundamental frequency Fig. 5 shows the equivalent circuit of the AC system for harmonic low order components where $S = j\omega_n$,

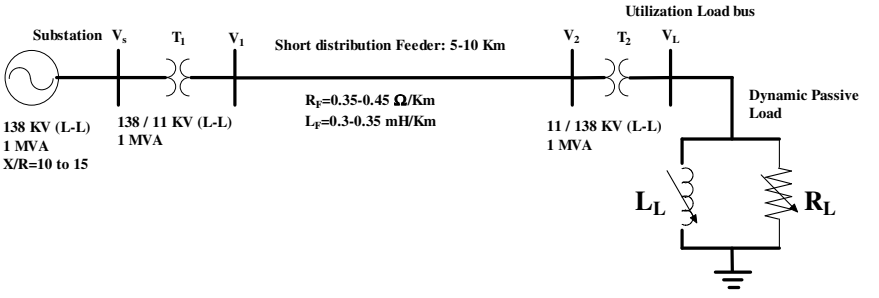


Fig. 17. Single Line Diagram for the Sample Radial Study Distribution AC System

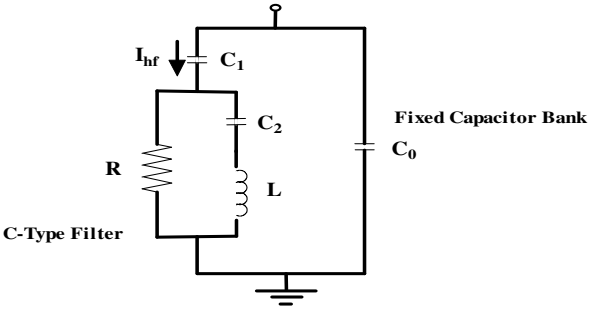


Fig. 18. Equivalent circuit of the hybrid power filter Compensator Scheme at the load bus

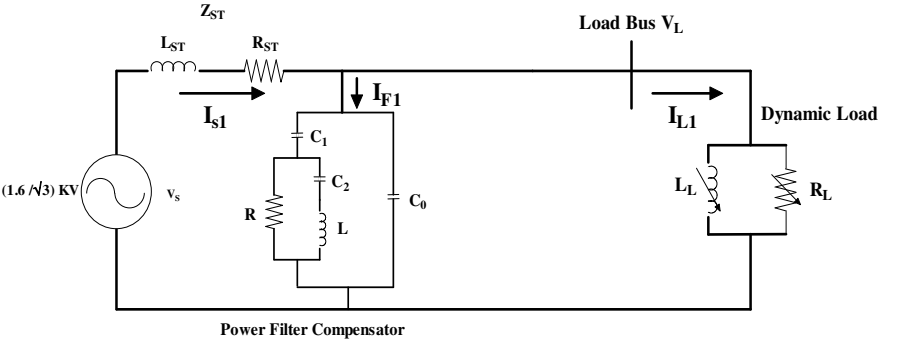


Fig. 19. Per Phase Equivalent AC circuit of the Sample Study System at the fundamental frequency component

n=3,5,7,9,11,13,15. The selection of the best and cost effective hybrid power filter parameters is very complicated due to the fact that the AC grid system's dynamic Thevenin's equivalent (driving-point) impedance is usually variable due to AC network switching, load variations, energization/de-energization of existing capacitor banks, static var compensators and shunt reactors. Harmonic AC equivalent system Thevenin's impedance can usually be measured or estimated

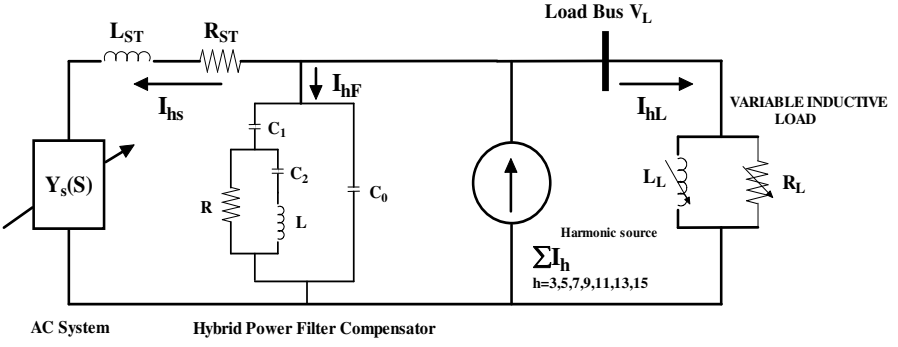


Fig. 20. Equivalent Per Phase AC circuit for n-harmonic frequency components

using Network/System Spectrum Analyzers and frequency domain identification software tools to obtain an effective equivalent system admittance as a curve-fitted continuous or discrete equivalent frequency model. Measurement over three phases may be required with an equivalent arithmetic or geometric averaging impedance model in the case of unbalanced phase conditions. The particle swarm optimization can handle any parametric variations in the AC system impedance with changes in the resultant optimization objective function. The power filter optimization problem including the key filter parametric constraints can be structured as follows:

(1) Calculate, measure and curve-fit the equivalent AC system equivalent harmonic impedance/admittance Ys(s),

$$Y_s(s) = \frac{1}{Z_s(s)} = \frac{b_1 + b_2 s + b_3 s^2 + b_4 s^3 + b_5 s^4}{a_1 + a_2 s + a_3 s^2 + a_4 s^3 + a_5 s^4 + a_6 s^5} \tag{38}$$

with $S = j\omega_n$, n=3,5,7,9,11,13,15 as a function of frequency over the specified selected harmonic frequency range. Ys(s) were obtained from measurement as:

$$Y_s(s) = \frac{N(s)}{D(s)} \tag{39}$$

Where:

$$N(s) = 8.60555 \times 10^{10} + 520833.3s + 729166.9s^2 + 0.7s^3 + 0.49s^4$$

$$D(s) = 1.071007 \times 10^{11} + 4.467016 \times 10^{11}s + 2267354s^2 + 2833543s^3 + 2.736762s^4 + 1.8355s^5$$

In addition, calculate the total equivalent dynamic impedances of the transformer and feeder…, this impedance may be considered constant:

$$Z_{ST} = R_{ST} + SL_{ST}$$

$$where: R_{ST} = 0.1\ \Omega, and\ L_{ST} = 20\ mH \tag{40}$$

The total equivalent Thevenin's admittance of the system with transformers and feeders is:

$$Y_{st} = \frac{1}{\left(\frac{1}{Y_s}\right) + (R_{st} + sL_{st})} \tag{41}$$

(2) Select the hybrid C-type power filter compensator structure with the equivalent transfer function, $Y_F(s)$, with $S = j\omega_n$, $n=3,5,7,9,11,13,15$. For the C-type filter, $Y_F(s)$ is a function of the power filter resistance R, the inductance L, and the capacitor sizes C_1, $C_2$and C_0 (see Fig. 3). $Y_F(s)$ represents the C-type filter selected and the compensated capacitor C_0:

$$Y_{Fn}(s) = \frac{sC_1(s^2LC_2 + RC_2s + 1)}{s^2LC_2(1+R) + RC_2s + R + 1} + SC_0 \tag{42}$$

(3) Calculate the dynamic variable equivalent load admittance $Y_L(S)$:

$$Y_{Ln}(S) = \frac{R_L + SL_L}{R_L \times SL_L} \tag{43}$$

$where:\ \ 0.5\ H < L_L < 2\ H$

$and \qquad 1\ \Omega < R_L < 5\ \Omega$

$S = j\omega_n$, $n=3,5,7,9,11,13,15.$

(4) Select the set of the discrete low order dominant or offending harmonics $S = j\omega_n$, $n=3,5,7,9,11,13,15$ over which the power filter optimization is to be performed using a multi objective function to reflect the basic power filter requirements. The main objective functions to be optimized with the following objectives are:

1. Minimize AC feeder current for AC system capacity release by reducing AC fundamental frequency feeder current,

$$Minimize\ J_1|_{S=j\omega_1} = |I_{S1}| \tag{44}$$

$$where: I_{S1} = \frac{\left(\frac{1.6 \times 10^3}{\sqrt{3}}\right)}{[R_{ST} + sL_{ST}] + \left(\frac{1}{Y_{F1} + Y_{L1}}\right)} \tag{45}$$

2. Minimize the voltage variations and change in fundamental frequency load bus voltage,

$$Minimize\ J_2|_{S=j\omega_1} = |\Delta V_{L1}| \tag{46}$$

$$where: |\Delta V_{L1}| = |I_{S1} \times [R_{ST} + sL_{ST}]| \tag{47}$$

3. Minimize the fundamental frequency RMS current squared to reduce active and reactive feeder power losses (See Fig. 4),

$$Minimize\ J_3|_{S=j\omega_1} = |I_{S1}|^2 \tag{48}$$

4. Minimize the harmonic current injection into the host electric grid system,

$$Minimize\ J_2|_{S=j\omega_h} = |G_{sh}(s)| \tag{49}$$

$$where: G_{sh}(s) = \frac{I_{hs}}{I_h} = \left|\frac{Y_s(s)}{Y_F(s) + Y_{st}(s) + Y_L(s)}\right| \tag{50}$$

5. Maximize the harmonic current absorption by the hybrid power filter compensator,

$$Maximize\ J_3|_{S=j\omega_h} = |G_{Fh}(s)| \tag{51}$$

$$where: G_{Fh}(s) = \frac{I_{hF}}{I_h} = \left|\frac{Y_F(s)}{Y_F(s) + Y_{st}(s) + Y_L(s)}\right| \tag{52}$$

6. Minimize any dynamic harmonic voltage distortion at the point of common coupling,

$$Minimize\ J_6|_{S=j\omega_h} = |G_{Vh}(s)| \tag{53}$$

$$where: G_{Vh}(s) = \frac{V_h}{I_h} = \left|\frac{1}{Y_F(s) + Y_{st}(s) + Y_L(s)}\right| \tag{54}$$

7. Define all hybrid power filter and capacitor compensator values and parameter constraints. For the C-type power filter with compensator unknown parameters {R, L, C_1, C_2, C_0}, the constraints are:

   (i) $R_{min} \le R \le R_{max}$

   (ii) $L_{min} \le R \le L_{max}$

   (iii) $|Y_s(s_k) + Y_F(s_k)| \ge \delta$, k = 1, 2..... m, where $\delta > 0$ is chosen so as to avoid any possible near-parallel resonance conditions on the combined power filter compensator with the system.

   (iv) either $C_i^{min} \le C_i \le C_i^{max}$, i = 1, 2, or C_i is a member of a discrete set {C_{i1}, C_{i2}, … , C_{ipi} }, i = 1, 2, of specified values which reflect other system requirements such as the need for reactive compensation or power factor correction.

8. Run iteratively the Multi objective Particle Swarm Optimization MOPSO routine to optimize the selected objective functions, with respect to the power filter parameters, to find the set of acceptable trade-off optimal solutions, the optimization process over all specified discrete dominant low order frequencies $S = j\omega_n$, $n=3,5,7,9,11,13,15$, where m is the selected higher harmonic order. Objective functions can be selected in pairs.

4.3.2 Digital Simulation Results

Two sample hybrid power filter compensator designs are presented.

Case 1: Hybrid Power Filter Compensator with Continuous Parameters Constraints

In this case, the hybrid power filter compensator parameters {R, L, C_1, C_2, C_0} are considered as continuously constrained to lie between specified bounds. The selected numerical values ranges are as follows:

$200\Omega \leq R \leq 800\Omega$

$0.05\,H \leq L \leq 0.5\,H$

$10\ \mu F \leq C_1 \leq 80\ \mu F$

$20\ \mu F \leq C_2 \leq 160\ \mu F$

$50\ \mu F \leq C_0 \leq 200\ \mu F$

$\omega_n \in \{3,5,7,9,11,13,15\} \times 377\ rad/sec.$, Seven discrete dominant offending harmonics.

Table 5 depicts the optimality solutions and final Pareto front surface compromise hybrid power filter and compensator.

Table 5. Final Pareto front-based Selections of the Hybrid power filter Compensator optimally selected parameters for the continuous case for different conflicting functions

R (Ω)	L (H)	C_1 (μF)	C_2 (μF)	C_0 (μF)	J1	J2	J3×10^4	J4×10^{-3}	J5	J6
769.06	0.1819	29.74	28.3	198.1	259.84	132.49	6.7521	0.079	1.1379	1.672
409.56	0.4501	35.47	95.4	148	261.28	133.23	6.8272	0.0695	1.1016	1.47
597.42	0.3441	33.09	101.5	101.8	262.85	134.02	6.9092	0.0618	1.0778	1.3073
536.99	0.2745	47.19	151.1	141.5	266.43	135.85	7.0989	0.0503	1.0495	1.064
704.70	0.3434	54.41	144.1	174.1	256.79	130.94	6.5944	0.115	1.3907	2.4331
516.35	0.3267	73.88	149.4	183.9	265.44	135.35	7.0462	0.0529	1.0552	1.1191
496.85	0.0858	74.65	156.6	111.2	271.32	138.34	7.3616	0.0409	1.0319	0.8654
752.45	0.3197	16.64	135.6	135.5	268.57	136.94	7.2131	0.0435	1.0364	0.9198
470.43	0.4166	77.46	117.4	92.5	258.83	131.98	6.6997	0.0924	1.2054	1.954
395.38	0.1639	24.78	74	152	263.83	134.52	6.9609	0.0578	1.067	1.2229
430.09	0.4214	12.81	156.9	147.9	257.47	131.28	6.6295	0.1033	1.2799	2.1859
667.39	0.2625	37.92	114.8	109.3	260.12	132.64	6.7667	0.0774	1.1312	1.6372
526.99	0.334	50.99	143.4	157.8	270.52	137.94	7.3186	0.042	1.0337	0.8885
395.32	0.4002	22.12	138.6	149.9	268.21	136.76	7.1939	0.0465	1.0418	0.9833
362.25	0.4607	67.45	42	166.8	272.02	138.70	7.3998	0.0405	1.0313	0.8575

Case 2: Hybrid Power Filter Compensator with Discrete Parameter Constraints

In this case, the resistance, the inductance, and the capacitor sizes C_1 and C_2 are chosen from a discrete set of specified parameter values. The discrete constraints of the hybrid power filter compensator parameters are as follows:

$$R \in \{200,300,400,500,600,700,800\}\ \Omega$$

$$L \in \{0.05,0.1,0.15,0.2,0.25,0.3,0.35,0.4,0.45.0.5\}\ H$$

$$\omega_n \in \{3,5,7,9,11,13\} \times 377\ rad/sec.$$

$$C_1 \in \{10,20,30,40,50,60,70,80\}\ \mu F$$

$$C_2 \in \{20,40,60,80,100,120,140,160\}\ \mu F$$

$$C_0 \in \{50,100,150,200,250\}\ \mu F$$

Table 6 shows the optimality selected possible solutions and Pareto selected parameters for the discrete MOPSO search.

Table 6. Final Pareto front-based Selections of the Hybrid power filter Compensator optimally selected parameters for the discrete case for different conflicting functions

R (Ω)	L (H)	C_1 (μF)	C_2 (μF)	C_0 (μF)	J1	J2	J3×10^4	J4×10^{-3}	J5	J6
300	0.4	60	140	200	271.57	138.47	7.3754	0.0405	1.0312	0.8566
800	0.3	50	100	150	265.45	135.35	7.0466	0.0528	1.055	1.117
500	0.5	80	180	250	278.87	142.2	7.7772	0.0328	1.0201	0.693
700	0.2	30	80	100	260.52	132.84	6.7873	0.0746	1.1201	1.5781
400	0.05	10	20	50	256.46	130.77	6.5776	0.1176	1.4212	2.4872

4.4 Micro Hydro-Fuel Cell Green Energy Management Scheme for Hydrogen and Island Electricity Generation

This application presents a novel FACTS based Electric Energy Management (EEM) compensator scheme based on Multi Objective Particle Swarm Optimization search technique MOPSO for use in hydrogen and Island Electricity Generation. It combines a fuel cell power source and a micro hydro water turbine. The novel control strategy is designed to achieve the high-efficiency coordinated operation of the two individual power sources and to regulate current and voltage for maximum utilization, without compromising the power quality and performance of the overall system. To achieve these conflicting objectives, a novel dual action Modulated Power Filter and Compensator at the AC bus (MPFC) and an additional Green Power Filter GPF scheme at the common DC bus using a dynamic self regulating error driven scheme for voltage stability, energy conservation, loss reduction, power factor correction, and power quality enhancement for hybrid multi source energy utilization systems. A tri-loop error driven dynamic controller is used to adjust the Pulse Width Modulation PWM switching of the DFC - Dynamic filter compensator on the AC side and green power filter on the DC side. Power factor correction and power quality enhancement is validated by simulation under different operating conditions, including sudden load disturbances and wind velocity excursions. Multi Objective Optimization MOPSO technique is used to find the optimal control gain settings that dynamically minimize the global dynamic error.

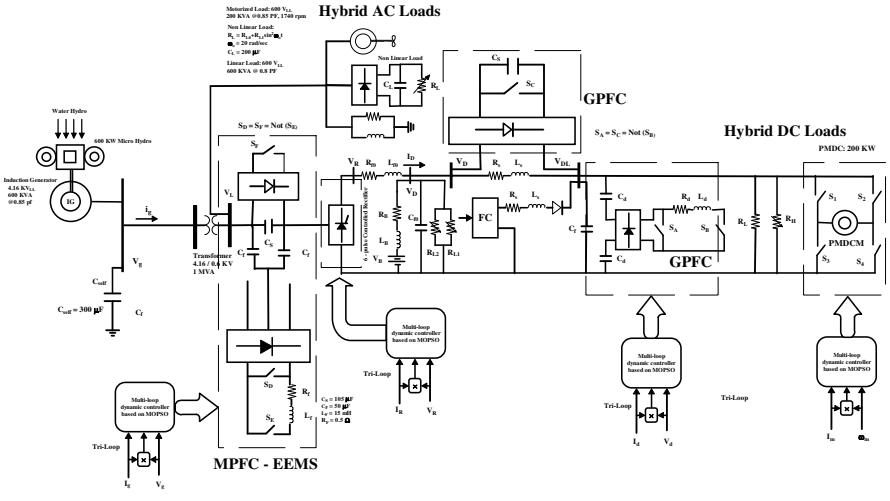


Fig. 21. Sample Study AC system with the novel AC Dynamic Filter Compensator (DFC) and the DC-side Green plug filter compensator

4.4.1 Sample Study System Description

The proposed study system comprises the following main components, as shown in Fig. 21. Micro hydro energy system, Fuel cell energy system, Battery bank backup system, Induction generator driven by the hydro turbine, Electrical energy management control system, Stabilization interface scheme and stabilization controller, The hybrid AC side load, and The hybrid DC side load.

Figures (22-29) depict the two dynamic self regulating PID and self tuned variable structure sliding mode electrical energy management dynamic controllers developed by the First Author for adjusting the switching duty-cycle-ratio based on Single Objective Particle Swarm Optimization SOPSO and Multi Objective Particle Swam Optimization MOPSO search and optimization techniques. The tri-loop error-driven dynamic controller is a novel structure developed by the First Author and used to control GPFC, MPFC, α- controller for the novel AC side converter and PMDCM. The global error is the summation of the three loop individual errors including voltage stability, current limiting and synthesize dynamic power loops. The global error signal is input to the self tuned PID controller or the self tuned variable structure sliding mode dynamic controller. As shown in figure (22) and (26), the (per-unit) three dimensional-error vector of the electrical energy management EEMS AC filter scheme is governed by the following equations:

$$e_{vg}(k) = V_{gref}(k) - V_g(k)\left(\frac{1}{1+ST_g}\right) \tag{55}$$

$$e_{Ig}(k) = I_g(k)\left(\frac{1}{1+ST_g}\right)\left(\frac{1}{1+SD}\right) - I_g(k)\left(\frac{1}{1+ST_g}\right) \tag{56}$$

$$e_{Pg}(k) = I_{gRMS}(k)\times V_{gRMS}(k)\left(\frac{1}{1+ST_g}\right)\left(\frac{1}{1+SD}\right) - I_{gRMS}(k)\times V_{gRMS}(k)\left(\frac{1}{1+ST_g}\right) \tag{57}$$

$$e_{PFg}(k) = PF_g(k)\left(\frac{1}{1+ST_g}\right)\left(\frac{1}{1+SD}\right) - PF_g(k)\left(\frac{1}{1+ST_g}\right) \tag{58}$$

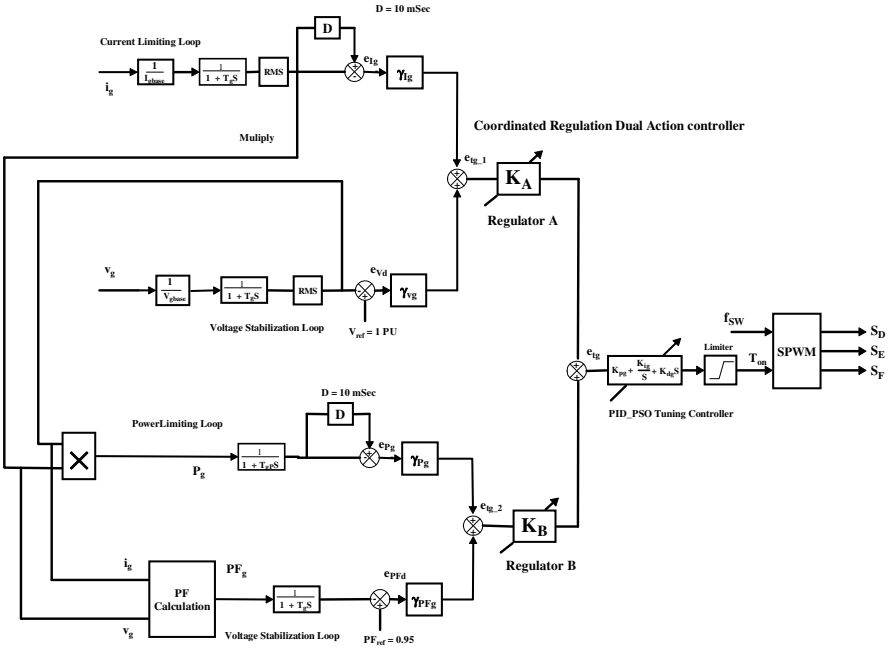


Fig. 22. Tri-loop error driven PSO self regulating PID dynamic controller for the AC side EEMS Scheme

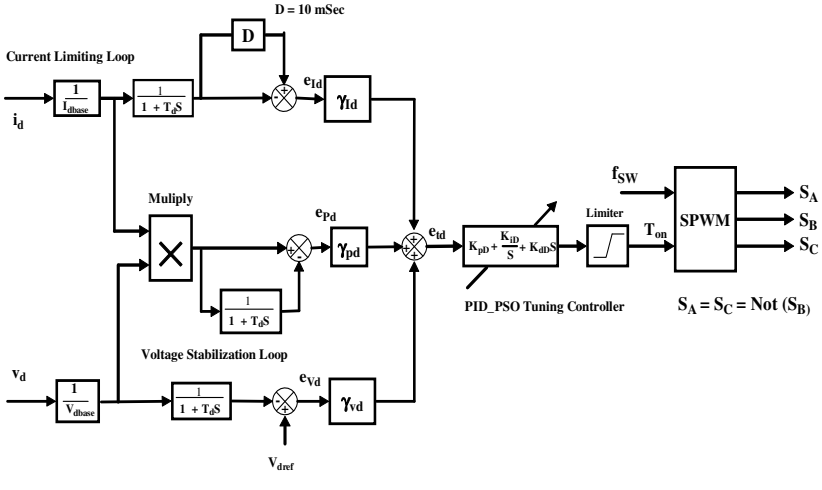


Fig. 23. Tri-loop error driven PSO self regulating PID dynamic controller for the DC side GPFC Scheme

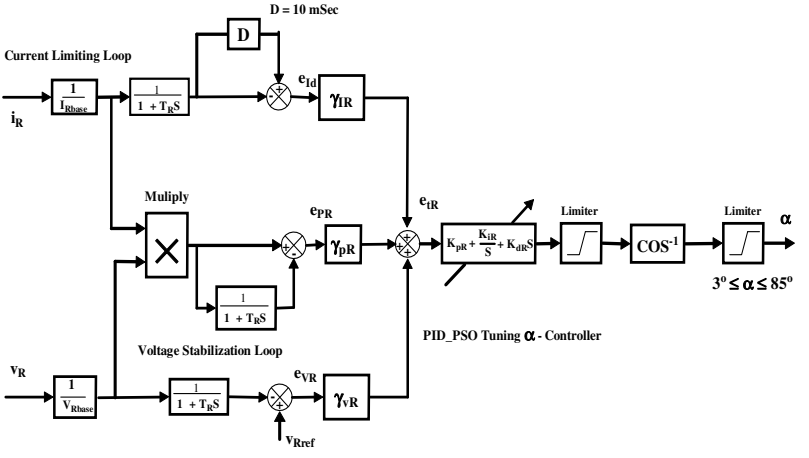


Fig. 24. Tri-loop error driven self regulating PID α- controller for the novel AC side converter- scheme

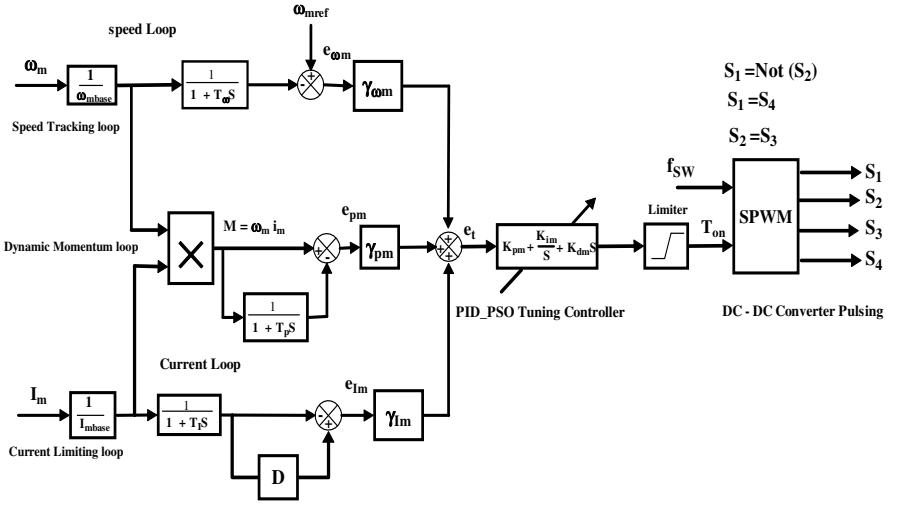


Fig. 25. Tri-loop error driven PSO self regulating PID dynamic controller for the speed control PMDC motor drive

The total or global error $e_{tg}(k)$ for the EEMS AC side scheme at a time instant:

$$e_{tg\_1}(k) = \gamma_{vg} e_{vg}(k) + \gamma_{ig} e_{ig}(k) \tag{59}$$

$$e_{tg\_2}(k) = \gamma_{Pg} e_{Pg}(k) + \gamma_{PFg} e_{PFg}(k) \tag{60}$$

$$e_{tg}(k) = K_A e_{tg\_1}(k) + K_B e_{tg\_2}(k) \tag{61}$$

Where: K_A, K_B are of the adjustable coordinated regulation weightings for the dual action controller parameters.

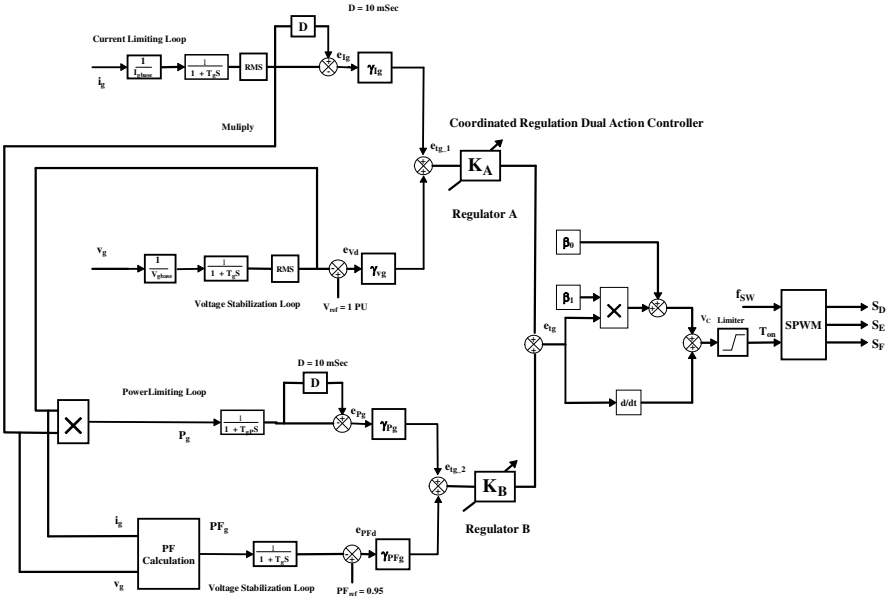


Fig. 26. Tri-loop error driven self tuned variable structure sliding mode dynamic controller the AC side EEMS filter scheme

The system control voltage has the following form in the time domain for the PID controller:

$$V_{cg}(t) = K_{pg} e_{tg}(t) + K_{ig} \int_0^t e_{tg}(t)dt + K_{dg} \frac{de_{tg}(t)}{dt} \tag{62}$$

While it has the following form in the time domain for self tuned variable structure sliding mode dynamic controller:

$$V_{cg}(t) = \beta_{0g} + \beta_{1g} e_{tg}(t) + \frac{de_{tg}(t)}{dt} \tag{63}$$

The on-line dynamic PSO optimization search algorithm is implemented for tuning the weightings (K_A, K_B) and the gains (K_{Pg}, K_{Ig}, K_{Dg}) or (β_{0g}, β_{1g}) to optimize the selected objective functions. In the same manner, figures (23) and (27) show the tri-loop error driven PSO self regulating PID dynamic and the self tuned variable structure sliding mode dynamic controllers of the GPFC DC-Side scheme. The (per-unit) three dimensional-error vector (e_{vd} ,e_{Id}, e_{pd}) is governed by the following equations:

$$e_{vd}(k) = V_{dref}(k) - V_d(k)\left(\frac{1}{1+ST_d}\right) \tag{64}$$

$$e_{Id}(k) = I_d(k)\left(\frac{1}{1+ST_d}\right)\left(\frac{1}{1+SD}\right) - I_d(k)\left(\frac{1}{1+ST_d}\right) \tag{65}$$

$$e_{Pd}(k) = I_d(k) \times V_d(k)\left(\frac{1}{1+ST_d}\right)\left(\frac{1}{1+SD}\right) - I_d(k) \times V_d(k)\left(\frac{1}{1+ST_d}\right) \tag{66}$$

And the total or global error e_{tg} (k) for the DC side green plug filter compensator GPFC scheme at a time instant:

$$e_{td}(k) = \gamma_{vd} e_{vd}(k) + \gamma_{id} e_{id}(k) + \gamma_{pd} e_{pd}(k) \tag{67}$$

The system control voltage of the GPFC scheme has the following form in the time domain for the PID controller:

$$V_{cd}(t) = K_{pd} e_{td}(t) + K_{id} \int_0^t e_{td}(t)dt + K_{dd} \frac{de_{td}(t)}{dt} \tag{68}$$

And the following time domain form for the novel self tuned variable structure sliding mode dynamic controller:

$$V_{cd}(t) = \beta_{0d} + \beta_{1d} e_{td}(t) + \frac{de_{td}(t)}{dt} \tag{69}$$

The PSO optimization algorithm is implemented for tuning the gains (K_{Pd}, K_{Id}, K_{Dg}) or the sliding surface parameters (β_{0d}, β_{1d}) to minimize the motor drive system total error e_{td}. In addition, figures (24) and (28) show the tri-loop error driven PSO self regulating PID dynamic and the self tuned variable structure sliding mode dynamic controllers of the of the three phase controlled rectifier scheme. The (per-unit) three dimensional-error vector (e_{vR} ,e_{IR}, e_{pR}) is governed by the following equations:

$$e_{vR}(k) = V_{Rref}(k) - V_R(k)\left(\frac{1}{1+ST_R}\right) \tag{70}$$

$$e_{IR}(k) = I_R(k)\left(\frac{1}{1+ST_R}\right)\left(\frac{1}{1+SD}\right) - I_R(k)\left(\frac{1}{1+ST_R}\right) \tag{71}$$

$$e_{PR}(k) = I_R(k) \times V_R(k)\left(\frac{1}{1+ST_R}\right)\left(\frac{1}{1+SD}\right) - I_R(k) \times V_R(k)\left(\frac{1}{1+ST_R}\right) \tag{72}$$

The total or global error e_{tg} (k) for the three phase controlled converter rectifier scheme at a time instant:

$$e_{tR}(k) = \gamma_{vR} e_{vR}(k) + \gamma_{iR} e_{iR}(k) + \gamma_{pR} e_{pR}(k) \tag{73}$$

The system control voltage of the three phase controlled rectifier scheme has the following form in the time domain for the PID controller:

$$V_{cR}(t) = K_{pR} e_{tR}(t) + K_{iR} \int_0^t e_{tR}(t)dt + K_{dR} \frac{de_{tR}(t)}{dt} \tag{74}$$

And the following time domain form for self tuned variable structure sliding mode dynamic controller:

$$V_{cR}(t) = \beta_{0R} + \beta_{1R} e_{tR}(t) + \frac{de_{tR}(t)}{dt} \tag{75}$$

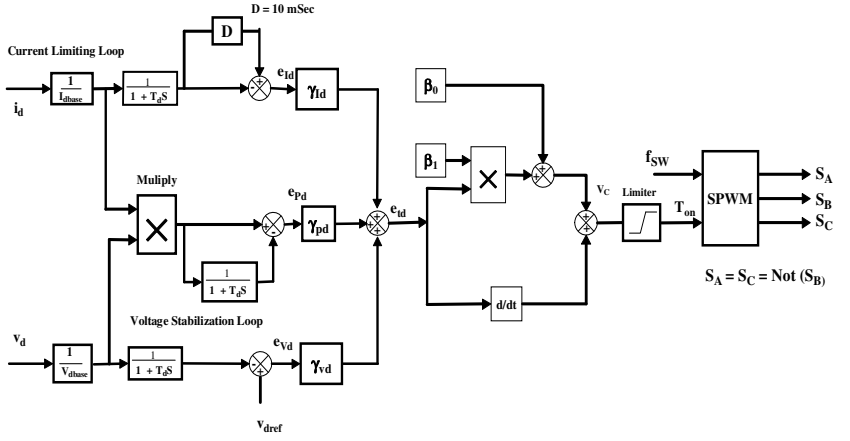


Fig. 27. Tri-loop error driven self tuned variable structure sliding mode dynamic controller for the DC side GPFC scheme

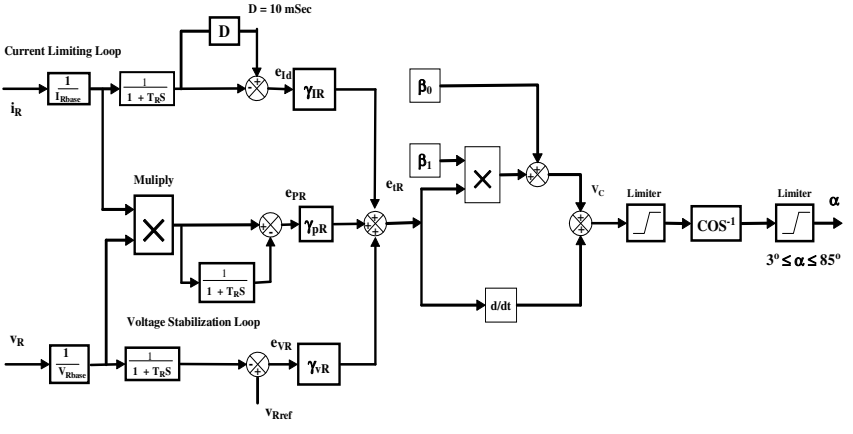


Fig. 28. Tri-loop error driven self tuned variable structure sliding mode dynamic α - controller for the AC side converter scheme

The first PSO search algorithm is implemented for on-line tuning of the gains (K_{PR}, K_{IR}, K_{DR}) or the sliding surface parameters (β_{0R}, β_{1R}) to minimize the selected objective functions. Finally, figures (25) and (29) show the tri-loop error driven PSO self regulating PID dynamic and the self tuned variable structure sliding mode dynamic controllers of the PMDC motor scheme. The (per-unit) three dimensional-error vector ($e_{\omega m}$,e_{Im}, e_{pm}) is governed by the following equations:

$$e_{\omega m}(k) = \omega_{mref}(k) - \omega_m(k)\left(\frac{1}{1 + ST_m}\right) \tag{76}$$

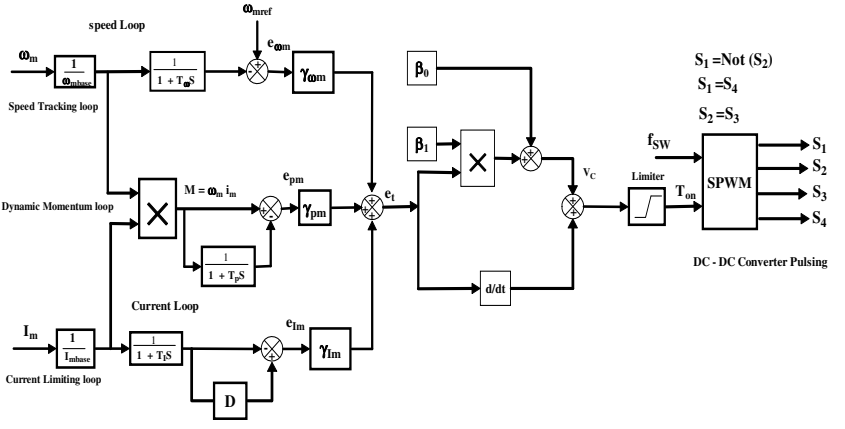


Fig. 29. Tri-loop error driven self tuned variable structure sliding mode dynamic speed controller for the PMDC motor drive

$$e_{\mathrm{Im}}(k) = I_m(k)\left(\frac{1}{1+ST_m}\right)\left(\frac{1}{1+SD}\right) - I_m(k)\left(\frac{1}{1+ST_m}\right) \tag{77}$$

$$e_{Pm}(k) = I_m(k)\times\omega_m(k)\left(\frac{1}{1+ST_m}\right)\left(\frac{1}{1+SD}\right) - I_m(k)\times\omega_m(k)\left(\frac{1}{1+ST_m}\right) \tag{78}$$

And the total or global error e_{tg} (k) for the EEMS scheme at a time instant:

$$e_{tm}(k) = \gamma_{\omega m} e_{\omega m}(k) + \gamma_{im} e_{im}(k) + \gamma_{pm} e_{pm}(k) \tag{79}$$

And the system control voltage has the following form in the time domain for the PID controller:

$$V_{cm}(t) = K_{pm} e_{tm}(t) + K_{im}\int_0^t e_{tm}(t)dt + K_{dm}\frac{de_{tm}(t)}{dt} \tag{80}$$

And the following time domain form for self tuned variable structure sliding mode dynamic controller:

$$V_{cm}(t) = \beta_{0m} + \beta_{1m} e_{tm}(t) + \frac{de_{tm}(t)}{dt} \tag{81}$$

The second PSO search algorithm is implemented for the dynamic on-line tuning of the gains (K_{Pm}, K_{Im}, K_{Dm}) or the other optimal sliding surface parameters (β_{0m}, β_{1m}) to minimize the selected objective functions. A number of conflicting objective functions are selected to optimize using the PSO algorithm. These functions are defined by the following:

1. **Minimize the voltage deviations at the AC and DC collection bus:**

$$Minimize \quad J_1 = \{\Delta V_{AC}, \Delta V_{AC}\} \tag{82}$$

$$\Delta V_{AC} = \left| V_g - V_L \right| \tag{83}$$

$$\Delta V_{DC} = \left| V_R - V_{DL} \right| \tag{84}$$

2. Minimize the grid collection Feeder total active power Losses:

$$J_2 = P_{P\_loss} = P_{P\_lossAC} + P_{P\_lossDC} \tag{85}$$

3. Minimize the grid collection Feeder total reactive power losses:

$$J_3 = P_{Q\_loss} = P_{Q\_lossAC} \tag{86}$$

4. Minimize the absolute total control error deviations:

$$J_4 = Minimize \quad \left\{ \left| e_{tg} \right|, \left| e_{tg} \right|, \left| e_{td} \right|, \left| e_{tm} \right| \right\} \tag{87}$$

4.4.2 Digital Simulation Results

SOPSO obtains a single global or near optimal solution based on a single selected weighted objective function. The weighted single objective function combines several objective functions using specified or selected weighting factors as follows:

$$weighted\ objective\ function = \alpha_1 J_1 + \alpha_2 J_2 + \alpha_3 J_3 + \alpha_4 J_4 \tag{88}$$

Where $\alpha_1 = 0.25$, $\alpha_2 = 0.25$, $\alpha_3 = 0.25$, $\alpha_4 = 0.25$, are selected weighting factors.

J_1 : Minimize the voltage deviations, J_2 : Minimize the grid collection feeder total active power Losses, J_3 : Minimize the grid collection feeder total reactive power Losses, and J_4 : Minimize the absolute control total error deviations. On the other hand, the MOPSO finds the set of acceptable (trade-off) Optimal Solutions. This set of accepted solutions is called Pareto front. These acceptable trade-off multi level solutions give more ability to the user to make an informed decision by seeing a wide range of near optimal selected solutions. The digital simulation results are presented for EEMS on the test system using MOPSO searching optimization algorithm. Table (7) shows the optimal solutions of the main objective functions versus the Tuned PID controller Gains based SOPSO and MOPSO control schemes and system behavior comparison without and with the EEMS Scheme based SOPSO and MOPSO optimization technique for the PID controller is shown in table (8). In addition to that, table (9) shows the optimal solutions of the main objective functions versus the self tuned variable structure sliding mode dynamic controller gains based SOPSO and MOPSO control scheme. Finally, Table (10) shows the system behavior comparison without and with the EEMS Scheme based SOPSO and MOPSO optimization technique for the self tuned variable structure sliding mode dynamic controller. Comparing the dynamic response results of the two study cases, without and with the Electrical Energy Management System EEMS, it is quite apparent that the Electrical Energy Management System EEMS highly improved the ac system dynamic performance from a general power quality point of view. The

effect of the EEMS is noticeable where it highly improved the power factor by reducing the amount of reactive power drawn from the supply, maintaining the system dynamic stability under sudden disturbances and recovering faster from any inrush transient state. The EEMS had a great impact on the supply power factor improving it from 0.223 to around 0.959 which is highly desired. Moreover, the Total Harmonic Distortion THD also improved from 25.46 % to a value of 4.85 %. In addition, the phase voltage can maintained around 1pu and the transient over-voltages and surge type inrush currents are also damped. Finally, the total losses are reduced from 0.1154 PU to a value of 0.0578 PU. By comparing the results of all sample simulation cases, it is concluded that the novel Electrical Energy Management System EEMS is an attractive low cost and efficient voltage stabilization and power factor correction device that also improve power quality and efficient-utilization of three phase –four wire Residential and Commercial Loads.

Table 7. Selected Objective Functions versus the Tuned PID controller Gains based SOPSO and MOPSO control schemes

	PID controller for the MPFC-AC side $\begin{bmatrix} K_{Pg} \\ K_{Ig} \\ K_{Dg} \end{bmatrix}$	PID controller for the GPFC – DC side $\begin{bmatrix} K_{Pd} \\ K_{Id} \\ K_{Dd} \end{bmatrix}$	PID controller for the PMDC motor $\begin{bmatrix} K_{Pm} \\ K_{\mathrm{Im}} \\ K_{Dm} \end{bmatrix}$	PID controller for α-controller for the novel AC side converter $\begin{bmatrix} K_{PR} \\ K_{IR} \\ K_{DR} \end{bmatrix}$	Coordinated Regulation Dual Action Controller Parameters $\begin{bmatrix} K_A \\ K_B \end{bmatrix}$	J_1 Minimize the voltage deviations (PU)	J_2 Minimize the Distribution Feeder total active power Losses (PU)	J_3 Minimize the Distribution Feeder total reactive power losses (PU)	J_4 Minimize the absolute total error deviations
SOPSO	$\begin{bmatrix} 71.4999 \\ 7.4489 \\ 1.2546 \end{bmatrix}$	$\begin{bmatrix} 29.1304 \\ 5.6018 \\ 2.2732 \end{bmatrix}$	$\begin{bmatrix} 85.5314 \\ 7.9876 \\ 1.0232 \end{bmatrix}$	$\begin{bmatrix} 66.5906 \\ 5.4041 \\ 0.7027 \end{bmatrix}$	$\begin{bmatrix} 0.71318 \\ 0.4415 \end{bmatrix}$	0.0650	0.0877	0.0856	0.2400
MOPSO	$\begin{bmatrix} 22.0395 \\ 2.6731 \\ 2.6278 \end{bmatrix}$	$\begin{bmatrix} 28.6419 \\ 7.3057 \\ 1.9051 \end{bmatrix}$	$\begin{bmatrix} 64.6479 \\ 9.8444 \\ 1.2982 \end{bmatrix}$	$\begin{bmatrix} 66.6899 \\ 8.2597 \\ 1.4372 \end{bmatrix}$	$\begin{bmatrix} 0.8486 \\ 0.5525 \end{bmatrix}$	0.0546	0.0726	0.0774	0.2090
	$\begin{bmatrix} 43.3429 \\ 7.3321 \\ 3.0695 \end{bmatrix}$	$\begin{bmatrix} 61.7633 \\ 5.3647 \\ 0.6842 \end{bmatrix}$	$\begin{bmatrix} 50.6282 \\ 2.0315 \\ 0.8048 \end{bmatrix}$	$\begin{bmatrix} 13.9506 \\ 6.9837 \\ 1.8053 \end{bmatrix}$	$\begin{bmatrix} 0.7385 \\ 0.4860 \end{bmatrix}$	0.0420	0.0426	0.0971	0.1823
	$\begin{bmatrix} 12.4467 \\ 4.2884 \\ 0.7927 \end{bmatrix}$	$\begin{bmatrix} 38.1417 \\ 2.2604 \\ 1.4376 \end{bmatrix}$	$\begin{bmatrix} 11.1576 \\ 6.1010 \\ 1.2220 \end{bmatrix}$	$\begin{bmatrix} 44.5571 \\ 8.4071 \\ 2.8921 \end{bmatrix}$	$\begin{bmatrix} 0.3742 \\ 0.2707 \end{bmatrix}$	0.0451	0.0515	0.0665	0.1774
	$\begin{bmatrix} 71.4804 \\ 7.0655 \\ 1.8065 \end{bmatrix}$	$\begin{bmatrix} 18.3558 \\ 9.0950 \\ 0.6783 \end{bmatrix}$	$\begin{bmatrix} 33.1804 \\ 5.6547 \\ 2.7004 \end{bmatrix}$	$\begin{bmatrix} 65.1156 \\ 1.5298 \\ 1.6746 \end{bmatrix}$	$\begin{bmatrix} 0.2741 \\ 0.7140 \end{bmatrix}$	0.0505	0.0734	0.0852	0.1615
	$\begin{bmatrix} 64.7686 \\ 4.2428 \\ 2.0176 \end{bmatrix}$	$\begin{bmatrix} 21.4184 \\ 6.7358 \\ 1.1605 \end{bmatrix}$	$\begin{bmatrix} 38.2770 \\ 5.9366 \\ 1.7460 \end{bmatrix}$	$\begin{bmatrix} 27.1067 \\ 3.3559 \\ 2.1819 \end{bmatrix}$	$\begin{bmatrix} 0.3725 \\ 0.5875 \end{bmatrix}$	0.0350	0.0466	0.0972	0.1511
	$\begin{bmatrix} 62.8227 \\ 6.3761 \\ 0.8274 \end{bmatrix}$	$\begin{bmatrix} 15.1823 \\ 4.2726 \\ 2.5418 \end{bmatrix}$	$\begin{bmatrix} 43.0811 \\ 1.4435 \\ 0.7385 \end{bmatrix}$	$\begin{bmatrix} 66.8306 \\ 7.3077 \\ 0.5434 \end{bmatrix}$	$\begin{bmatrix} 0.2358 \\ 0.7281 \end{bmatrix}$	0.0256	0.0534	0.0791	0.2683

Table 8. System dynamic behavior comparison with and without the EEMS Scheme based SOPSO and MOPSO optimization technique for the Proportional – Integral – Derivative PID controller

	without the EEMS Scheme	With the EEMS Scheme with constant controller gains	with the EEMS Scheme with SOPSO Optimization Technique	with the EEMS Scheme with MOPSO Optimization Technique
RMS Voltage (PU)	0.82896	0.9119	0.9604	0.9843
Power Factor	0.27093	0.8243	0.9228	0.9705
THD (%) Total Harmonic Distorsion	23.9662	9.0278	4.8802	3.7656
Maximum Transient Voltage – Over/Under Shoot (PU)	0.1426	0.0958	0.0779	0.0576
Maximum Transient Current – Over/Under Shoot (PU)	0.0996	0.0608	0.0538	0.0355
RMS Current (PU)	0.6931	0.5592	0.4905	0.4223
Active Power Losses (PU)	0.0970	0.0610	0.0483	0.0379
Reactive Power Losses (PU)	0.1104	0.0879	0.0613	0.0516

Table 9. Selected Objective Functions versus the self tuned variable structure sliding mode dynamic controller gains based SOPSO and MOPSO control scheme

	self tuned variable structure sliding mode dynamic controller for the MPFC-AC side $\begin{bmatrix}\beta_{og}\\ \beta_{1g}\end{bmatrix}$	self tuned variable structure sliding mode dynamic controller for the GPFC – DC side $\begin{bmatrix}\beta_{od}\\ \beta_{1d}\end{bmatrix}$	self tuned variable structure sliding mode dynamic controller for the PMDC motor $\begin{bmatrix}\beta_{om}\\ \beta_{1m}\end{bmatrix}$	self tuned variable structure sliding mode dynamic controller for α-controller for the novel AC side converter $\begin{bmatrix}\beta_{oR}\\ \beta_{1R}\end{bmatrix}$	Coordinated Regulation Dual Action Controller Parameters $\begin{bmatrix}K_A\\ K_B\end{bmatrix}$	J_1 Minimize the voltage deviations (PU)	J_2 Minimize the Distribution Feeder total active power Losses (PU)	J_3 Minimize the Distribution Feeder total reactive power losses (PU)	J_4 Minimize the absolute total error deviations
SOPSO	$\begin{bmatrix}0.8660\\ 22.214\end{bmatrix}$	$\begin{bmatrix}1.7330\\ 20.245\end{bmatrix}$	$\begin{bmatrix}0.8948\\ 9.6204\end{bmatrix}$	$\begin{bmatrix}1.4895\\ 13.7441\end{bmatrix}$	$\begin{bmatrix}0.3450\\ 0.2789\end{bmatrix}$	0.0628	0.0808	0.0940	0.2817
MOPSO	$\begin{bmatrix}1.4032\\ 17.963\end{bmatrix}$	$\begin{bmatrix}0.7750\\ 32.968\end{bmatrix}$	$\begin{bmatrix}1.5288\\ 24.726\end{bmatrix}$	$\begin{bmatrix}0.5193\\ 15.052\end{bmatrix}$	$\begin{bmatrix}0.1137\\ 0.7721\end{bmatrix}$	0.0399	0.0865	0.0875	0.2029
	$\begin{bmatrix}0.9656\\ 18.4001\end{bmatrix}$	$\begin{bmatrix}1.6686\\ 36.173\end{bmatrix}$	$\begin{bmatrix}0.6115\\ 29.794\end{bmatrix}$	$\begin{bmatrix}1.2744\\ 16.8786\end{bmatrix}$	$\begin{bmatrix}0.5006\\ 0.9386\end{bmatrix}$	0.0487	0.0555	0.0690	0.2583
	$\begin{bmatrix}1.1873\\ 17.0769\end{bmatrix}$	$\begin{bmatrix}1.3737\\ 33.6440\end{bmatrix}$	$\begin{bmatrix}0.7899\\ 30.1731\end{bmatrix}$	$\begin{bmatrix}0.9147\\ 26.2635\end{bmatrix}$	$\begin{bmatrix}0.5194\\ 0.4768\end{bmatrix}$	0.0572	0.0534	0.0863	0.1734

Table 9. (*Continued*)

	$\begin{bmatrix}1.6563\\25.7307\end{bmatrix}$	$\begin{bmatrix}0.9709\\18.9640\end{bmatrix}$	$\begin{bmatrix}1.2543\\10.3980\end{bmatrix}$	$\begin{bmatrix}1.9216\\23.7090\end{bmatrix}$	$\begin{bmatrix}0.8616\\0.5726\end{bmatrix}$	0.0519	0.0668	0.0581	0.1970
	$\begin{bmatrix}1.7420\\22.1173\end{bmatrix}$	$\begin{bmatrix}1.8763\\25.6158\end{bmatrix}$	$\begin{bmatrix}0.6696\\21.3020\end{bmatrix}$	$\begin{bmatrix}1.7182\\33.674\end{bmatrix}$	$\begin{bmatrix}0.2824\\0.7049\end{bmatrix}$	0.0326	0.0482	0.0978	0.2036
	$\begin{bmatrix}0.7346\\27.675\end{bmatrix}$	$\begin{bmatrix}1.898\\10.436\end{bmatrix}$	$\begin{bmatrix}0.6832\\42.564\end{bmatrix}$	$\begin{bmatrix}1.6440\\18.645\end{bmatrix}$	$\begin{bmatrix}0.8543\\0.1177\end{bmatrix}$	0.0426	0.0505	0.0598	0.1541

Table 10. System dynamic behavior comparison with and without the EEMS Scheme based SOPSO and MOPSO optimization technique for the self tuned Variable Structure Sliding Mode Dynamic Controller

	without the EEMS Scheme	With the EEMS Scheme with constant controller gains	with the EEMS Scheme with SOPSO Optimization Technique	with the EEMS Scheme with MOPSO Optimization Technique
RMS Voltage (PU)	0.82896	0.9246	0.9402	0.9878
Power Factor	0.27093	0.8755	0.9422	0.9646
THD (%) Total Harmonic Distortion	23.9662	7.9994	6.6523	5.0405
Maximum Transient Voltage – Over/Under Shoot (PU)	0.1426	0.0617	0.0474	0.0382
Maximum Transient Current – Over/Under Shoot (PU)	0.0996	0.0688	0.0400	0.0308
RMS Current (PU)	0.6931	0.5240	0.4712	0.4307
Active Power Losses (PU)	0.0970	0.0770	0.0593	0.0474
Reactive Power Losses (PU)	0.1104	0.0834	0.0605	0.0574

4.5 An Optimally Coordinated Wind-FC-Diesel Utilization Scheme

The modeling and coordinated control strategy for a hybrid wind–diesel-fuel cell renewable green energy generation system with battery backup is investigated in this application. The proposed hybrid scheme tracks the maximum power efficiency and optimal energy capture from the wind, diesel and the fuel cell. The diesel generator set unit issued to balance the system demand power according to the total power demand in order to minimize the diesel fuel consumption. The wind system generator is regulated to maximize the energy capture from the wind turbine. The paper presents a novel application of Multi Objective Particle Swarm Optimization MOPSO technique to control the 6-pulse rectifier converter,

dynamic filter/capacitor compensation DFC and Green Power Filter GPF schemes developed by the First Author using real time dynamic self regulating error tracking. A tri-loop error driven dynamic controller is used to adjust the PWM switching of the DFC- Dynamic AC side filter compensator and green power filter. Full power factor correction and power quality enhancement is validated under different operating conditions, including sudden load disturbances and wind velocity excursions. Multi Objective Optimization MOPSO technique is used to find the near optimal control gain settings that dynamically minimize the global dynamic error for each controller.

4.5.1 Sample Study System Description

A Standalone Wind Energy Conversion Scheme (WECS) using induction generator (IG) is studied in this paper under a time sequence of Load variations, Excursions and Wind gusting. The standalone WECS connected to the local load bus over a radial transmission line. The WECS comprises the following main components, as shown in Fig. 30: Wind turbine, Gear box, Synchronous Generator driven by the Diesel Engine, Fuel Cell, Battery Backup System, Induction generator driven by the wind turbine, Stabilization interface scheme and stabilization controller, and the hybrid electric load. All the parameters of the unified AC study system of the standalone WECS sample study system are given in Appendix. The sample standalone WECS was subjected to severe combined sequence of load switching/load variation/load excursion and wind speed variation and gusting.

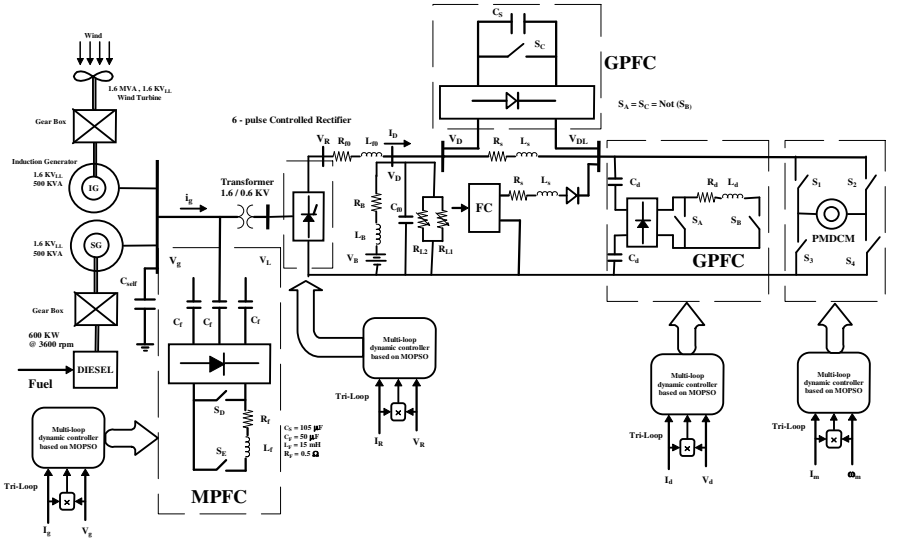


Fig. 30. Sample three-phase Study ac system with the novel AC Dynamic Filter Compensator (DFC) and DC-side Green plug filter compensator

A. Diesel Generator - Set

From an electrical system point of view, a diesel generator can be represented as a prime mover and generator. Ideally, the prime mover has the capability to supply any power demand up to rated power at constant frequency. The synchronous generator connected to it must be able to keep the voltage constant at any load condition. The diesel engine kept the frequency constant by maintaining the rotor speed. When power demand fluctuates the diesel generator could vary its output via fuel regulation to its governor. The synchronous generator must control its output voltage by controlling the excitation current. Thus, the diesel generating system, as a unit, must be able to control its frequency and its output voltage. The ability of the diesel generator to respond to frequency changes was affected by the inertia of the diesel gen-set, the sensitivity of the governor, and the power capability of the diesel engine. The ability of the synchronous generator to control its voltage was affected by the field-winding time constant, the availability of the DC power to supply the field winding, and the response of voltage control regulation.

B. Wind-Turbine Generator

The power generated by the wind turbine was defined as follows:

$$T_W = \frac{1}{2\lambda}\rho A R C_P V_W^2 = \frac{1}{2\omega_W}\rho A C_P V_W^3 = k\frac{V_W^3}{\omega_W} \tag{89}$$

Where

ρ is the specified density of air (1.25kg/m^2)

A is the area swept by the blades

R is the radius of the rotor blades

C_P is power conversion coefficient

λ is the tip speed ratio

ω_W is the wind turbine velocity in rpm

k is equivalent coefficient in per unit (0.745)

The novel sinusoidal PWM Switched Power Filter schemes using real time dynamic voltage, current and power as tri-loop error tracking inputs to the proposed Particle swarm optimization controller (POS) scheme are shown in Fig. (6). The AC side dynamic filter and compensator scheme can also be an attractive solution for both distribution and utilization radial circuits, feeding a nonlinear type load. Figures (31-38) depict the dynamic self regulating PID controller and self tuned variable structure sliding mode dynamic controller for adjusting the switching duty-cycle-ratio based on Multi Objective Particle swam Optimization search and optimization technique MOPSO. The effective reactance of the combined hybrid fixed capacitors and the modulated tuned arm filter depends on the duty cycle and the frequency of the SPWM output which in turn is a function of the self tuned variable structure sliding mode controller output. The output of

the SPWM generator is a train of pulses with variable duty cycles and constant frequency. The degree of filtering and compensation is dependent on the duty cycle of the generated pulses. This would in turn vary the effective reactance of the hybrid power filter. The tri-loop error-driven dynamic controller is a novel structure developed by the First Author and used to control GPFC, MPFC, and PMDCM. The global error is the summation of the three loop individual errors

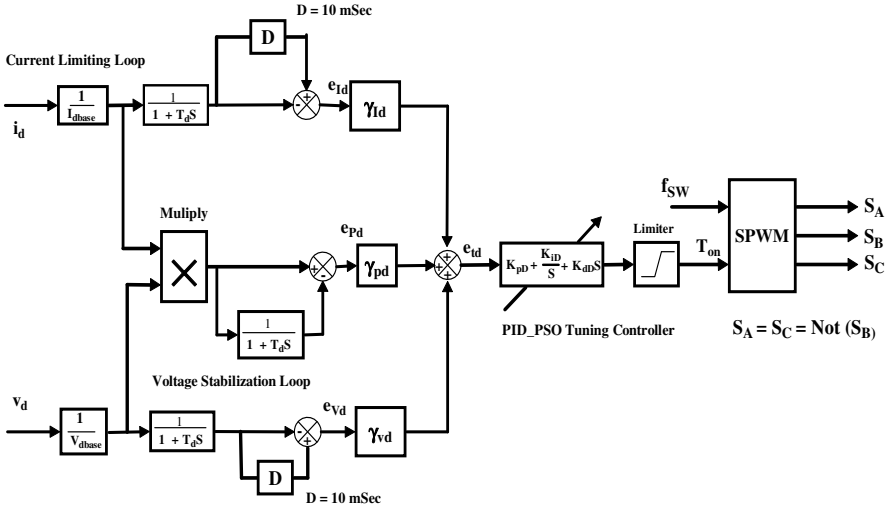


Fig. 31. Tri-loop error driven PSO self regulating PID dynamic controller for the DC side GPFC Scheme

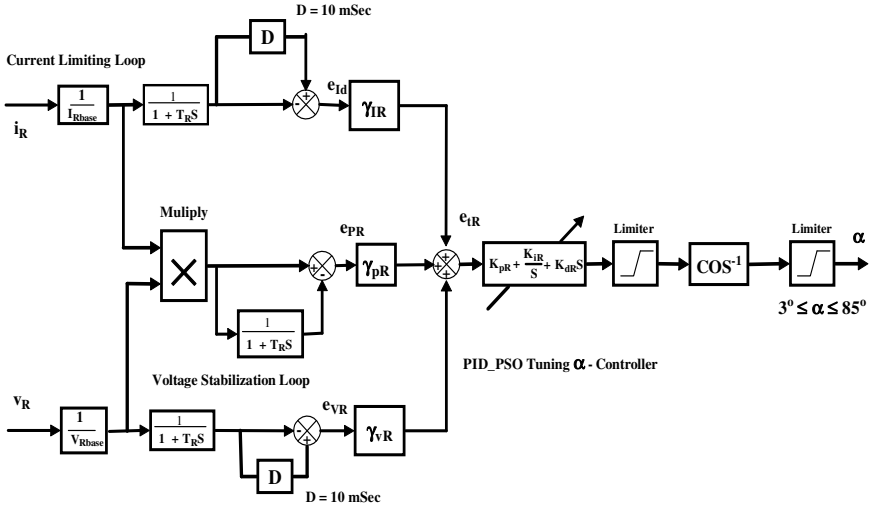


Fig. 32. Tri-loop error driven self regulating PID α- controller for the novel AC side converter- scheme

including voltage stability, current limiting and synthesize dynamic power loops. The global error signal is input to the self tuned PID controller. The (per-unit) three dimensional-error vector (e_{vg} ,e_{Ig}, e_{pg}) of the MPFC AC filter scheme is governed by the following equations:

$$e_{vg}(k) = V_g(k)\left(\frac{1}{1+ST_g}\right)\left(\frac{1}{1+SD}\right) - V_g(k)\left(\frac{1}{1+ST_g}\right) \tag{90}$$

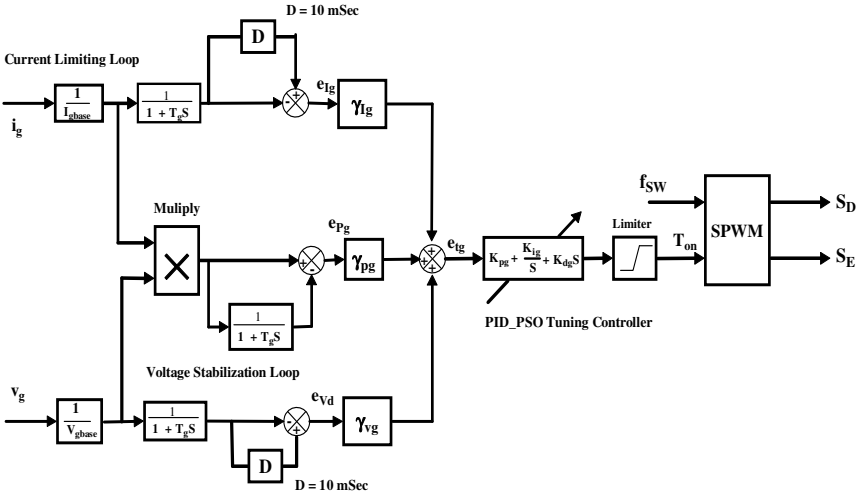


Fig. 33. Tri-loop error driven PSO self regulating PID dynamic controller for the AC side MPFC Scheme

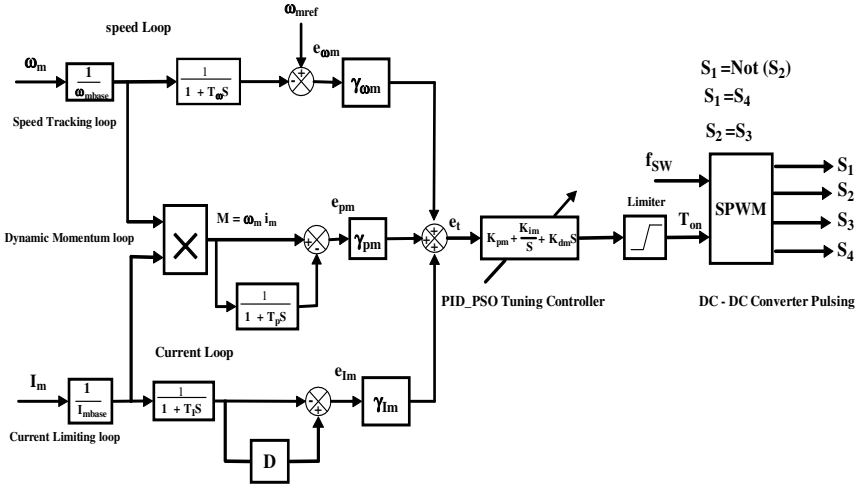


Fig. 34. Tri-loop error driven PSO self regulating PID dynamic controller for the speed control PMDC motor drive

$$e_{Ig}(k) = I_g(k)\left(\frac{1}{1+ST_g}\right)\left(\frac{1}{1+SD}\right) - I_g(k)\left(\frac{1}{1+ST_g}\right) \tag{91}$$

$$e_{Pg}(k) = I_g(k)\times V_g(k)\left(\frac{1}{1+ST_g}\right)\left(\frac{1}{1+SD}\right) - I_g(k)\times V_g(k)\left(\frac{1}{1+ST_g}\right) \tag{92}$$

The total or global error $e_{tg}(k)$ for the MPFC AC side scheme at a time instant:

$$e_{tg}(k) = \gamma_{vg}e_{vg}(k) + \gamma_{ig}e_{ig}(k) + \gamma_{pg}e_{pg}(k) \tag{93}$$

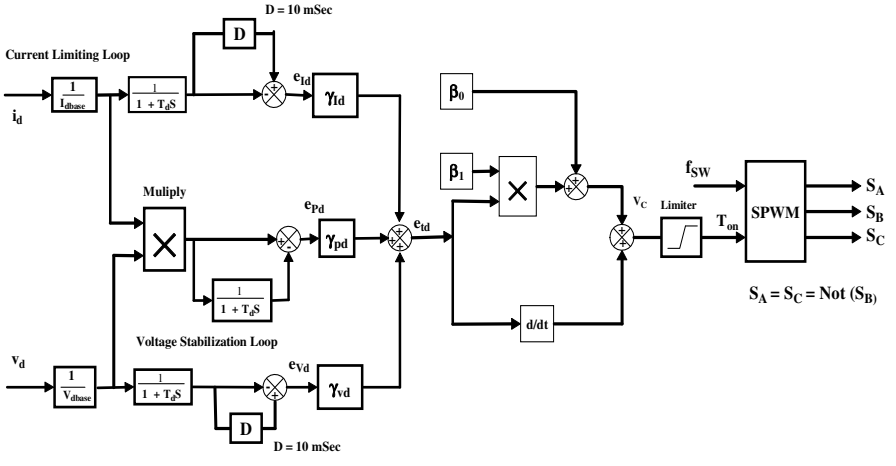


Fig. 35. Tri-loop error driven self tuned variable structure sliding mode dynamic controller for the DC side GPFC scheme

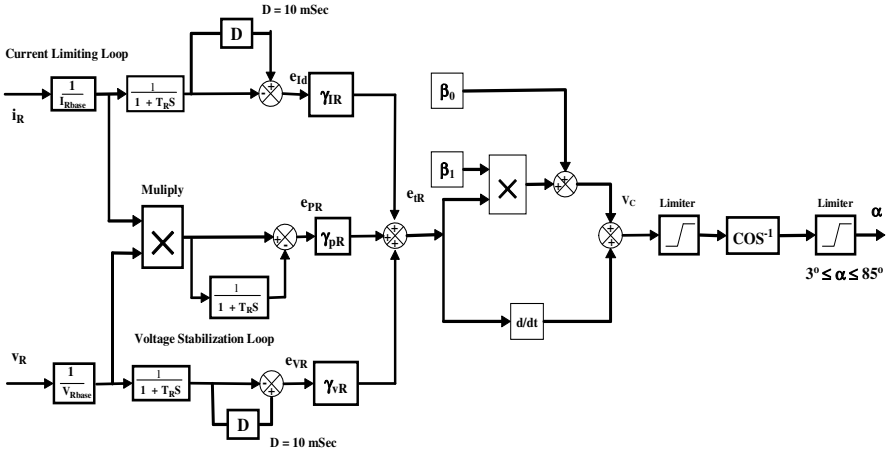


Fig. 36. Tri-loop error driven self tuned variable structure sliding mode dynamic α - controller for the AC side converter scheme

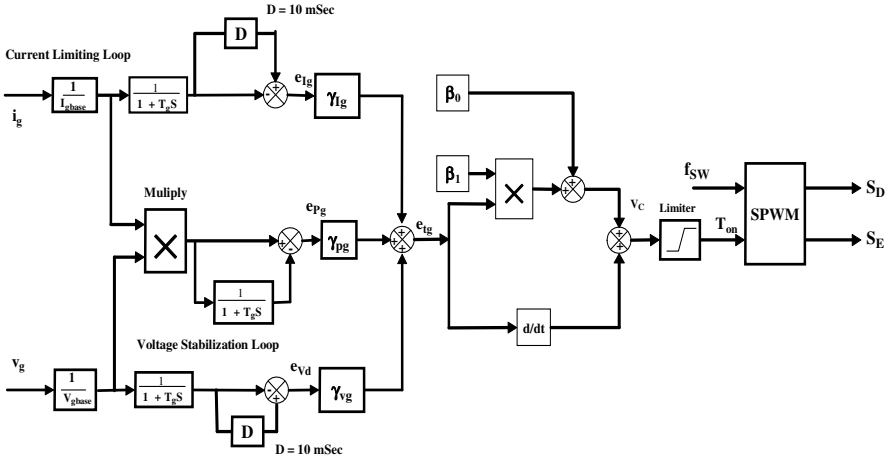


Fig. 37. Tri-loop error driven self tuned variable structure sliding mode dynamic controller the AC side MPFC filter scheme

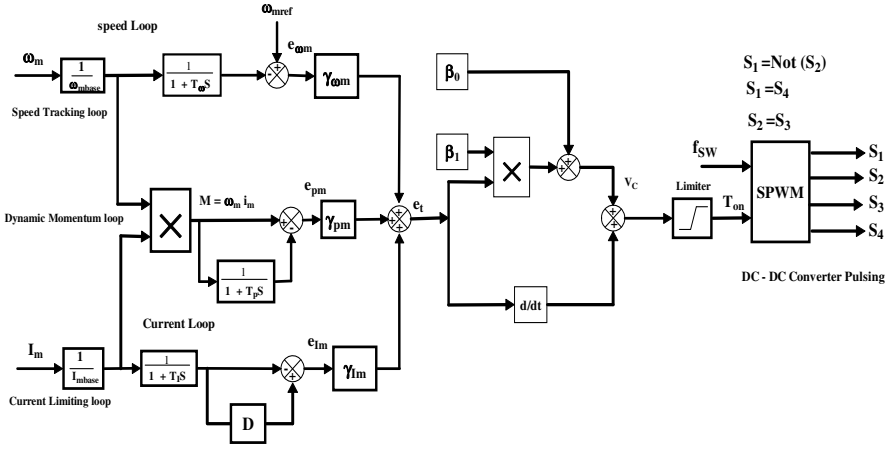


Fig. 38. Tri-loop error driven self tuned variable structure sliding mode dynamic speed controller for the PMDC motor drive

The system control voltage has the following form in the time domain for the PID controller:

$$V_{cg}(t) = K_{pg} e_{tg}(t) + K_{ig} \int_0^t e_{tg}(t)dt + K_{dg} \frac{de_{tg}(t)}{dt} \tag{94}$$

While it has the following form in the time domain for self tuned variable structure sliding mode dynamic controller:

$$V_{cg}(t) = \beta_{0g} + \beta_{1g} e_{tg}(t) + \frac{de_{tg}(t)}{dt} \tag{95}$$

The PSO optimization search algorithm is implemented for tuning the gains (K_{Pg}, K_{Ig}, K_{Dg}) or (β_{0g}, β_{1g}) to minimize the system total error e_{tg}. In the same manner, The (per-unit) three dimensional-error vector (e_{vd} ,e_{Id}, e_{pd}) of the GPFC scheme is governed by the following equations:

$$e_{vd}(k) = V_d(k)\left(\frac{1}{1+ST_d}\right)\left(\frac{1}{1+SD}\right) - V_d(k)\left(\frac{1}{1+ST_d}\right) \tag{96}$$

$$e_{Id}(k) = I_d(k)\left(\frac{1}{1+ST_d}\right)\left(\frac{1}{1+SD}\right) - I_d(k)\left(\frac{1}{1+ST_d}\right) \tag{97}$$

$$e_{Pd}(k) = I_d(k)\times V_d(k)\left(\frac{1}{1+ST_d}\right)\left(\frac{1}{1+SD}\right) - I_d(k)\times V_d(k)\left(\frac{1}{1+ST_d}\right) \tag{98}$$

And the total or global error e_{tg} (k) for the DC side green plug filter compensator GPFC scheme at a time instant:

$$e_{td}(k) = \gamma_{vd}\, e_{vd}(k) + \gamma_{id}\, e_{id}(k) + \gamma_{pd}\, e_{pd}(k) \tag{99}$$

The system control voltage of the GPFC scheme has the following form in the time domain for the PID controller:

$$V_{cd}(t) = K_{pd} e_{td}(t) + K_{id}\int_0^t e_{td}(t)dt + K_{dd}\frac{de_{td}(t)}{dt} \tag{100}$$

And the following time domain form for self tuned variable structure sliding mode dynamic controller:

$$V_{cd}(t) = \beta_{0d} + \beta_{1d} e_{td}(t) + \frac{de_{td}(t)}{dt} \tag{101}$$

The PSO optimization algorithm is implemented for tuning the gains (K_{Pd}, K_{Id}, K_{Dg}) or (β_{0d}, β_{1d}) to minimize the motor drive system total error e_{td}. In addition, The (per-unit) three dimensional-error vector (e_{vR} ,e_{IR}, e_{pR}) of the three phase controlled rectifier scheme is governed by the following equations:

$$e_{vR}(k) = V_R(k)\left(\frac{1}{1+ST_R}\right)\left(\frac{1}{1+SD}\right) - V_R(k)\left(\frac{1}{1+ST_R}\right) \tag{102}$$

$$e_{IR}(k) = I_R(k)\left(\frac{1}{1+ST_R}\right)\left(\frac{1}{1+SD}\right) - I_R(k)\left(\frac{1}{1+ST_R}\right) \tag{103}$$

$$e_{PR}(k) = I_R(k)\times V_R(k)\left(\frac{1}{1+ST_R}\right)\left(\frac{1}{1+SD}\right) - I_R(k)\times V_R(k)\left(\frac{1}{1+ST_R}\right) \tag{104}$$

The total or global error e_{tg} (k) for the three phase controlled converter rectifier scheme at a time instant:

$$e_{tR}(k) = \gamma_{vR}\, e_{vR}(k) + \gamma_{iR}\, e_{iR}(k) + \gamma_{pR}\, e_{pR}(k) \tag{105}$$

The system control voltage of the three phase controlled rectifier scheme has the following form in the time domain for the PID controller:

$$V_{cR}(t) = K_{pR} e_{tR}(t) + K_{iR} \int_0^t e_{tR}(t)dt + K_{dR} \frac{de_{tR}(t)}{dt} \tag{106}$$

And the following time domain form for self tuned variable structure sliding mode dynamic controller:

$$V_{cR}(t) = \beta_{0R} + \beta_{1R} e_{tR}(t) + \frac{de_{tR}(t)}{dt} \tag{107}$$

The PSO search and optimization algorithm is implemented for tuning the gains (K_{PR}, K_{IR}, K_{DR}) to minimize the motor drive system total error e_{tR}. or (β_{0R}, β_{1R}).

Finally, the (per-unit) three dimensional-error vector ($e_{\omega m}$,e_{Im}, e_{pm}) of the PMDC motor scheme is governed by the following equations:

$$e_{\omega m}(k) = \omega_m(k)\left(\frac{1}{1+ST_m}\right)\left(\frac{1}{1+SD}\right) - \omega_m(k)\left(\frac{1}{1+ST_m}\right) \tag{108}$$

$$e_{\mathrm{Im}}(k) = I_m(k)\left(\frac{1}{1+ST_m}\right)\left(\frac{1}{1+SD}\right) - I_m(k)\left(\frac{1}{1+ST_m}\right) \tag{109}$$

$$e_{Pm}(k) = I_m(k)\times\omega_m(k)\left(\frac{1}{1+ST_m}\right)\left(\frac{1}{1+SD}\right) - I_m(k)\times\omega_m(k)\left(\frac{1}{1+ST_m}\right) \tag{110}$$

And the total or global error e_{tg} (k) for the MPFC scheme at a time instant:

$$e_{tm}(k) = \gamma_{\omega m} e_{\omega m}(k) + \gamma_{im} e_{im}(k) + \gamma_{pm} e_{pm}(k) \tag{111}$$

And the system control voltage has the following form in the time domain for the PID controller:

$$V_{cm}(t) = K_{pm} e_{tm}(t) + K_{im} \int_0^t e_{tm}(t)dt + K_{dm} \frac{de_{tm}(t)}{dt} \tag{112}$$

And the following time domain form for self tuned variable structure sliding mode dynamic controller:

$$V_{cm}(t) = \beta_{0m} + \beta_{1m} e_{tm}(t) + \frac{de_{tm}(t)}{dt} \tag{113}$$

The PSO searching algorithm is implemented for tuning the gains (K_{Pm}, K_{Im}, K_{Dm}) or (β_{0m}, β_{1m}) to minimize the system total error e_{tm}.

4.5.2 Digital Simulation Results

The parallel operation of a hybrid wind turbine, with the diesel generator set, and the fuel cell performance is compared for the two cases, with fixed and self tuned type controller. Matlab-Simulink Software was used to design, test, and validate the effectiveness of the two FACTS devices and the associated dynamic SPWM controllers. Table (11) shows system behavior comparison with and without

MPFC&GPFC using PID and tuned variable structure sliding mode dynamic controllers, comparing the dynamic response results of the two study cases, with and without MPFC & GPFC. The simulation results show the system dynamic response for unbalanced hybrid nonlinear loading condition with constant and tuned controller parameters. In addition, the figures show the system dynamic response for three-phase short duration fault and open phase fault with fixed and self tuned controller parameters. The three phase short circuit fault is single line to ground fault (Phase A) from 0.2 to 0.4 sec at generator bus and the open phase fault is from 0.6 to 0.8 sec in phase A. Comparing the dynamic response results, with fixed and self tuned controller parameters, it is quite apparent that the dynamic filter compensator with tuned controller parameters did highly improve the ac system dynamic performance from a general power quality point of view. The effect of the AC side DFC filter is noticeable where it highly enhanced the power factor by reducing the amount of reactive power drawn from the generator supply, maintaining the AC system dynamic stability under sudden disturbances and recovering faster from any transient excursions. In addition, the phase voltage can be also maintained around 1pu and the transient over-voltages and surge type inrush currents are also damped. By comparing the results of all sample digital simulation cases, it is concluded that the two dynamic FACTS devices (DFC + GPFC) with the SPWM dynamic multi-loop controller developed by the First Author can be used as efficient Voltage stabilization and power factor correction tools to improve power quality and efficient energy utilization.

Table 11. System behavior comparison with and without MPFC&GPFC using PID and tuned variable structure sliding mode dynamic controllers

	without the MPFC and GPFC	with the MPFC and GPFC with constant parameters of PID dynamic controller	with the MPFC and GPFC with constant parameters of variable structure sliding mode dynamic controller	with the MPFC and GPFC with tuned PID dynamic controller	with the MPFC and GPFC with tuned variable structure sliding mode dynamic controller
RMS Voltage (PU)	0.7807	0.918467	0.902847	0.96744	0.97827
Power Factor	0.256534	0.896746	0.848756	0.9564	0.9690
THD (%) Total Harmonic Distortion	28.7341	9.45987	8.95982	3.995674	3.64582
Maximum Transient Voltage – Over/Under Shoot (PU)	0.25647	0.128546	0.140934	0.0453	0.0544
Maximum Transient Current – Over/Under Shoot (PU)	0.209645	0.129354	0.118534	0.0853	0.0751
RMS Current (PU)	0.8881	0.536745	0.516783	0.40893	0.39674
Active Power Losses (PU)	0.109345	0.092673	0.093564	0.0701	0.0509
Reactive Power Losses (PU)	0.1156743	0.10356	0.108474	0.0540	0.06103

References

[1] Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings, IEEE International Conf. on Neural Networks, vol. 4, pp. 1942–1948

[2] Shi, Y., Eberhart, R.: Empirical study of particle swarm optimization. In: Proceedings of the 1999 Congress on Evolutionary Computation, vol. 3 (1999)

[3] Shi, Y., Eberhart, R.: A modified particle swarm optimizer. In: IEEE International Conf. on Evolutionary Computation, pp. 69–73 (1998)
[4] Eberhart, R., Shi, Y.: Particle swarm optimization: developments, applications and resources. In: Proceedings of the 2001 Congress on Evolutionary Computation, vol. 1, pp. 81–86 (2001)
[5] Shi, Y., Eberhart, R.: Parameter Selection in Particle Swarm Optimization. In: Porto, V.W., Waagen, D. (eds.) EP 1998. LNCS, vol. 1447, pp. 591–601. Springer, Heidelberg (1998)
[6] Fukuyama, Y., et al.: A Particle Swarm Optimization for Reactive Power and Voltage Control Considering Voltage Security Assessment. IEEE Transaction on Power Systems 15(4) (November 2000)
[7] Gudise, V.G., Venayagamoorthy, G.K.: Evolving Digital Circuits Using Particle Swarm. In: Proceedings of the International Joint Conference on Neural Networks, July 20-24, vol. 1, pp. 468–472 (2003)
[8] Ngatchou, P., Zarei, A., El-Sharkawi, A.: Pareto Multi Objective Optimization. In: Proceedings of the 13th International Conference on Intelligent Systems Application to Power Systems, November 6-10, pp. 84–91 (2005)
[9] Berizzi, A., Innorta, M., Marannino, P.: Multiobjective optimization techniques applied to modern power systems. In: 2001 IEEE Power Engineering Society Winter Meeting, January 28-February 1 (2001)
[10] Coello Coello, C.A., Lechuga, M.S.: Mopso: A proposal for multiple objective particle swarm optimization. In: IEEE Proceedings World Congress on Computational Intelligence, pp. 1051–1056 (2003)
[11] Fieldsend, J.E., Singh, S.: A multi-objective algorithm based upon particle swarm optimisation, an efficient data structure and turbulence. In: The 2002 U.K. Workshop on Computational Intelligence, pp. 34–44 (2002)
[12] Hu, X., Eberhart, R., Shi, Y.: Particle swarm with extended memory for multiobjective optimization. In: IEEE Swarm Intelligence Symposium, pp. 193–198 (2003)
[13] Mostaghim, S., Teich, J.: Strategies for finding good local guides in multi-objective particle swarm optimization. In: IEEE Swarm Intelligence Symposium, pp. 26–33 (2003)
[14] Gaing, Z.-L.: Discrete Particle Swarm Optimization Algorithm for Unit Commitment. In: IEEE Power Engineering Society General Meeting, July 13-17, vol. 1 (2003)
[15] Kennedy, J., Eberhart, R.: A Discrete Binary Version of the Particle Swarm Optimization. In: Proceedings of IEEE International Conference on Neural Networks, Perth, Australia, vol. IV, pp. 4104–4108 (1997)
[16] del Valle, Y., Venayagamoorthy, G.K., Mohagheghi, S., Hernandez, J.-C., Harley, R.G.: Particle Swarm Optimization: Basic Concepts, Variants and Applications in Power Systems. IEEE Transactions on Evolutionary Computation 12(2), 171–195 (2008)
[17] Hamid, R.H., Abdel Amin, A.M.A., Ahmed, R.S., Elgammal, A.A.A.: Optimal Operation of Induction Motors Based on Multi-Objective Particle Swarm Optimization (MOPSO). In: The 33rd Annual Conference of the IEEE Industrial Electronics Society, Taipei, Taiwan, November 5-8 (2007)
[18] Sharaf, A.M., El-Gammal, A.A.A.: A Dynamic MOPSO Self Regulating Modulated Power Filter Compensator Scheme for Electric Distribution Networks. The Journal of the Association of Professional Engineers of Trinidad & Tobago, ISSN 1000 7924

[19] Sharaf, A.M., El-Gammal, A.A.A.: A Novel Discrete Multi-Objective Particle Swarm Optimization (MOPSO) Technique for Optimal Hybrid Power Filter Compensator Schemes. International Journal of Power and Energy Conversion 1(2/3), 157–177 (2009)

[20] Sharaf, A.M., El-Gammal, A.A.A.: Novel Particle Swarm Optimization PSO Self Regulating Control Schemes for Dynamic Error Driven PMDC Motor Drives. Accepted to publish in International Journal of Power and Energy Conversion

[21] Sharaf, A.M., El-Gammal, A.A.A.: A Dynamic MOPSO Self Regulating Modulated Power Filter Compensator Scheme for Electric Distribution Networks. The Journal of the Association of Professional Engineers of Trinidad & Tobago, ISSN 1000 7924

[22] Sharaf, A.M., El-Gammal, A.A.A.: A Novel Particle Swarm Optimization PSO-Based Dynamic Green Power Filter-Compensator for Radial Distribution Systems. Accepted to publish in International Journal of Power Engineering & Green Technology (IJPEGT)

[23] Sharaf, A.M., El-Gammal, A.A.A.: An Optimal Voltage and Energy Utilization for a Stand-Alone Wind Energy Conversion Scheme WECS Based on Particle Swarm Optimization PSO. Accepted to publish in International Journal of Power Engineering & Green Technology (IJPEGT)

[24] Sharaf, A.M., El-Gammal, A.A.A.: A Novel Discrete Multi-Objective Particle Swarm Optimization (MOPSO) Technique for Optimal Hybrid Power Filter Compensator Schemes. International Journal of Power and Energy Conversion 1(2/3), 157–177 (2009)

[25] Sharaf, A.M., El-Gammal, A.A.A.: A Multi Objective Multi-Stage Particle Swarm Optimization MOPSO Search Scheme for Power Quality and Loss Reduction on Radial Distribution System. RE&PQJ-7, Renewable Energy and Power Quality Journal (7) (April 2009)

[26] Sharaf, A.M., El-Gammal, A.A.A.: Optimal Design of Hybrid Power Filter Compensator Using Multi-Objective Particle Swarm Optimization (MOPSO). Al-Azhar University Engineering international scientific Journal (JAUES), ISSN 1110-6409

[27] El-Gammal, A.A.A., El-Samahy, A.A.: Optimal Tuning of PID Controller on a DC Motor Drives Using Particle Swarm Optimization PSO. Al-Azhar University Engineering international scientific Journal (JAUES), ISSN 1110-6409

[28] Sharaf, A.M., El-Gammal, A.A.A.: Optimal Switched Dynamic Modulated Power Filter Compensator for Radial Distribution System. In: 2009 Third UKSim European Symposium on Computer Modeling and Simulation, European Modeling Symposium, EMS 2009, Athens, Greece, November 25-27 (2009)

[29] Sharaf, A.M., El-Gammal, A.A.A.: A Particle Swarm Optimization Technique (PSO) for Power Filter Design. In: 2009 Third UKSim European Symposium on Computer Modeling and Simulation, European Modeling Symposium, EMS 2009, Athens, Greece, November 25-27 (2009)

[30] Sharaf, A.M., El-Gammal, A.A.A.: Optimal Multi-Incremental Self Regulating Speed Controller for Industrial PMDC Motor Drive Systems. In: 2009 Third UKSim European Symposium on Computer Modeling and Simulation, European Modeling Symposium, EMS 2009, Athens, Greece, November 25-27 (2009)

[31] Sharaf, A.M., El-Gammal, A.A.A.: A MOPSO Self Regulating PID Dynamic Error Driven Controller for Tidal Wave Energy Conversion. In: 2009 Third UKSim European Symposium on Computer Modeling and Simulation, European Modeling Symposium, EMS 2009, Athens, Greece, November 25-27 (2009)

[32] Sharaf, A.M., El-Gammal, A.A.A.: Optimal Self Regulating PID Controller for Coordinated Wind-FC-Diesel Utilization Scheme. In: 2009 Third UKSim European Symposium on Computer Modeling and Simulation, European Modeling Symposium, EMS 2009, Athens, Greece, November 25-27 (2009)

[33] Sharaf, A.M., El-Gammal, A.A.A.: A MOPSO Tri-loop Self Regulated Variable Structure Sliding Mode Self Regulating Coordinated Controller for Tidal Wave Energy Conversion. In: 2009 Third UKSim European Symposium on Computer Modeling and Simulation, European Modeling Symposium, EMS 2009, Athens, Greece, November 25-27 (2009)

[34] Sharaf, A.M., El-Gammal, A.A.A.: A MOPSO Tri-loop Self Regulated Variable Structure Sliding Mode Self Regulating Coordinated Controller for Tidal Wave Energy Conversion. In: 2009 Third UKSim European Symposium on Computer Modeling and Simulation, European Modeling Symposium, EMS 2009, Athens, Greece, November 25-27 (2009)

[35] Sharaf, A.M., El-Gammal, A.A.A.: Optimal Self Regulating Green Plug Compensator Scheme for Electric Distribution Networks. In: The International Conference on Electric Power and Energy Conversion Systems, EPECS 2009, November 10-12. American University of Sharjah, UAE (2009)

[36] Sharaf, A.M., El-Gammal, A.A.A.: Optimal Tan-Sigmoid Tuned Efficient Control Strategy for Industrial PMDC Motor Drive Systems. In: The International Conference on Electric Power and Energy Conversion Systems, EPECS 2009, November 10-12. American University of Sharjah, UAE (2009)

[37] Sharaf, A.M., El-Gammal, A.A.A.: Optimal Self Regulating Stand-Alone Wave Energy Conversion Scheme. In: Electrical Power and Energy Conference Sustainable / Renewable Energy Systems and Technologies, EPEC 2009, Montreal, Quebec, Canada, October 22-23 (2009)

[38] Sharaf, A.M., El-Gammal, A.A.A.: Optimal Energy Utilization for a Stand-Alone Wind Energy Scheme WES. In: Electrical Power and Energy Conference Sustainable / Renewable Energy Systems and Technologies, EPEC 2009, Montreal, Quebec, Canada, October 22-23 (2009)

[39] Sharaf, A.M., El-Gammal, A.A.A.: A Novel Error Driven Dynamic Tri-Loop Controller Based on Multi Stage Particle Swarm Optimization-MSPSO for Industrial PMDC Motor Drives. In: INISTA 2009, International Symposium on INnovations in Intelligent SysTems and Applications, Trabzon, Turkey, June 29-July 1 (2009)

[40] Sharaf, A.M., El-Gammal, A.A.A.: An Incremental Sliding mode Variable Structure Particle Swarm Optimization-PSO Optimal Tuned Controller for Industrial PMDC Motor Drives. In: The IEEE International Electric Machines and Drives Conference, IEMDC 2009, Miami, Florida, May 3-6 (2009)

[41] Sharaf, A.M., El-Gammal, A.A.A.: An Integral Squared Error -ISE Optimal Parameters Tuning of Modified PID Controller for Industrial PMDC Motor Based on Particle Swarm Optimization-PSO. In: The 2009 IEEE 6th International Power Electronics and Motion Control Conference - ECCE Asia, IPEMC 2009, Wuhan, China, May 17-20 (2009)

[42] Sharaf, A.M., El-Gammal, A.A.A.: Optimal Tuning of Tri-Loop Dynamic Error Driven Controller for Industrial PMDC Motor Drives Based on Particle Swarm Optimization-PSO. In: The 2009 IEEE 6th International Power Electronics and Motion Control Conference - ECCE Asia, IPEMC 2009, Wuhan, China, May 17-20 (2009)

[43] Sharaf, A.M., El-Gammal, A.A.A.: A Novel Particle Swarm Optimization PSO Tuning Scheme for PMDC Motor Drives Controllers. In: The Second International Conference on Power Engineering, Energy and Electrical Drives, POWERENG 2009 (2009)

[44] Sharaf, A.M., El-Gammal, A.A.A.: A Multi Objective Multi-Stage Particle Swarm Optimization MOPSO Search Scheme for Power Quality and Loss Reduction on Radial Distribution System. In: International Conference on Renewable Energies and Power Quality, ICREPQ 2009 (2009)

[45] El-Gammal, A.A.A., El-Samahy, A.A.: A Modified Design For Optimal PID Controller For DC Motor Drives Using Particle Swarm Optimization PSO. In: The Second International Conference on Power Engineering, Energy and Electrical Drives, POWERENG 2009 (2009)

[46] Sharaf, A.M., El-Gammal, A.A.A.: Optimal Hybrid Power Filter Compensator Design Using Multi-Objective Particle Swarm Optimization (MOPSO). In: 11th International Conference on Computer Modeling and Simulation, UKSIM 2009 (2009)

[47] El-Gammal, A.A.A., El-Samahy, A.A.: Adaptive Tuning of a PID Speed Controller For DC Motor Drives Using Multi-Objective Particle Swarm Optimization MOPSO. In: 11th International Conference on Computer Modeling and Simulation, UKSIM 2009 (2009)

[48] Sharaf, A.M., El-Gammal, A.A.A.: A Novel Discrete Multi-Objective Particle Swarm Optimization (MOPSO) of Optimal Shunt Power Filter. In: Power Systems Conference and Exposition, PSCE 2009, Seattle, Washington, USA (2009)

Particle Swarm Optimization and Its Applications in Power Systems

M.R. AlRashidi, M.F. AlHajri, A.K. Al-Othman, and K.M. El-Naggar

1 Introduction

Optimization problems are widely encountered in various fields in science and technology. The fact that most optimization problems, when modeled accurately, are of non-convex and sometimes discrete nature has encouraged many researchers to develop new optimization techniques to overcome such difficulties. Particle Swarm Optimization (PSO) is one of the newly developed optimization techniques with many attractive features. Early experimentations of employing PSO in many applications in science and technology have indicated its promising potential. Thus, the basics of PSO theory, development, main features, and its applications in power systems are presented in the following sections.

2 Fundamentals of Particle Swarm Optimization

Two scientists, namely Kennedy and Eberhart, first introduced PSO in 1995 as a new metaheuristic method [1;2]. They studied a stochastic nonlinear model that was developed by Heppner and Grenander to simulate species movement traveling in groups [3]. The original objective of the research conducted by Heppner and Grenander was to create a computer model that simulates the social behavior of bird flocks and fish schools. As Kennedy and Eberhart progressed in their

M.R. AlRashidi
College of Technological Studies (PAAET), Kuwait
mr.alrashidi@paaet.edu.kw

M.F. AlHajri
College of Technological Studies (PAAET), Kuwait
mf.alhajri@paaet.edu.kw

A.K. Al-Othman
College of Technological Studies (PAAET), Kuwait
ak.alothman@paaet.edu.kw

K.M. El-Naggar
College of Technological Studies (PAAET), Kuwait
km.elnaggar@paaet.edu.kw

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 295–324.
springerlink.com

research, they discovered that with some modifications the social behavior model can also serve as a powerful optimizer. They realized that such species try to approach their target in an optimal manner which resembles finding the optimal solution to any mathematical optimization problem. The first version of PSO was intended to handle only nonlinear continuous optimization problems.

One of the key attractive features of the PSO approach is its simplicity as it involves only two model equations. In PSO, the coordinates of each particle represent a possible solution associated with two vectors, the position (xi) and velocity (vi) vectors. In N-dimensional search space, $x_i = [x_{i1}, x_{i2}, \ldots x_{iN}]$ and $v_i = [v_{i1}, v_{i2}, \ldots v_{iN}]$ are the two vectors associated with each particle i. A swarm consists of a number of particles "or solution candidates" that fly through the feasible solution space to explore points where optimal solutions exist. During their search, particles interact with each other in a certain way as to optimize their search experience. In each iteration, the particle with the best solution shares its position coordinates (*Gbest*) information with the rest of the swarm. Then, each particle updates its coordinates based on its own best search experience (*Pbest*) as well as (*Gbest*) according to the following equations:

$$v_i^{k+1} = v_i^k + c_1 r_1 (Pbest_i^k - x_i^k) + c_2 r_2 (Gbest^k - x_i^k) \quad (1)$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \quad (2)$$

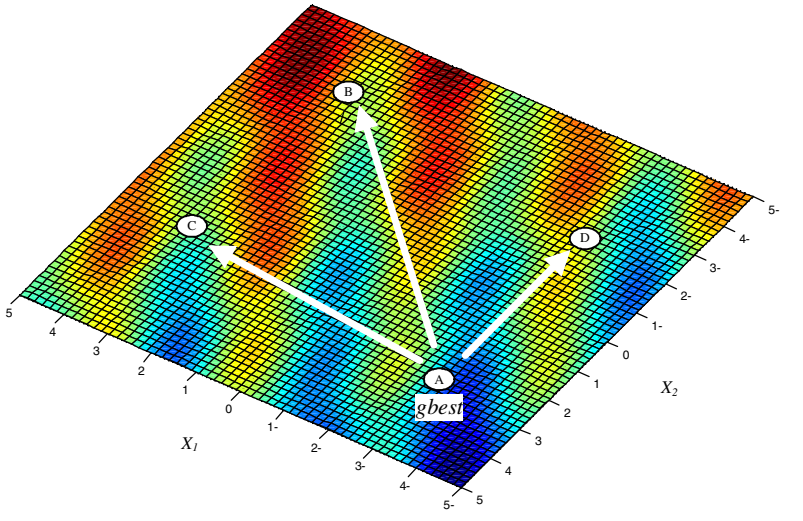
where

- c_1 and c_2 are two positive acceleration constants, they keep balance between the particle's individual and social behavior when they are set to be equal.
- r_1 and r_2 are two randomly generated numbers with a range of [0,1] added in the model to introduce stochastic nature to the particles' movement.
- $Pbest_i^k$ is the best position particle *i* achieved based on its own experience; $Pbest_i^k = [x_{i1}^{pbest}, x_{i2}^{pbest}, ..., x_{iN}^{pbest}]$
- $Gbest^k$ is the best particle position based on overall swarm's experience; $Gbest^k = [x_1^{gbest}, x_2^{gbest}, ..., x_N^{gbest}]$
- *k* is the iteration index

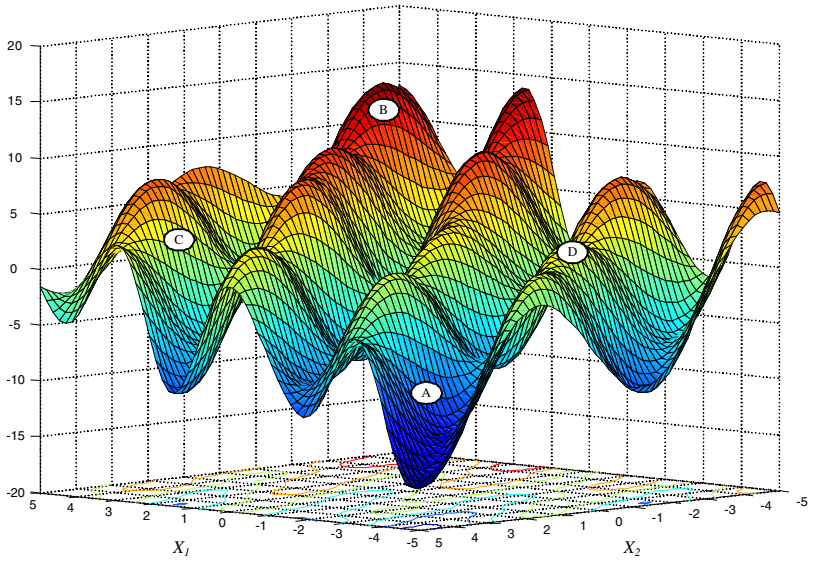
Equations (1) and (2) represent the original PSO model equations introduced in 1995. However, this model experienced poor convergence characteristics and sometimes additional fitness evaluations were needed to find an optimal solution.

3 PSO Development

Many advances in PSO development elevated its capabilities to handle a wider class of complex engineering and science optimization problems and to improve its overall performance. Summaries of recent advances in these areas are presented in references [4-6]. Different variants of the PSO algorithm were proposed but the most standard is the global version of PSO (*Gbest* model) introduced by Shi and



2D



3D

Fig. 1. Information sharing mechanism in *Gbest* model

Eberhart [7], in which the whole population is considered as a single neighborhood throughout the optimization process. Information sharing mechanism among swarm members of the *Gbest* model can be best described using graphical illustration. A mathematical function of two varibles with multiple valleys and peaks is depicted in Fig. 1. The known global solution is found to be (-4.7119, 4.7116) with minimum objective value of -16.4248. In the *Gbest* model, the entire swarm of four members, namely A, B, C, and D forms a single neighborhood. A snapshot of each particle location during their flying experience (in iteration k) is also shown in Fig. 1. It is clear that particle A is the closest one to the global solution (i.e. having minimum objective value) in this iteration. Thus, it will send its *Gbest* coordinates to the rest of the swarm members.

The original model Equations (1) and (2) are modified in the *Gbest* model as follows:

$$v_i^{k+1} = \underbrace{wv_i^k}_{\text{previous velocity}} + \underbrace{c_1 r_1 (Pbest_i^k - x_i^k)}_{\text{cognitive component}} + \underbrace{c_2 r_2 (Gbest^k - x_i^k)}_{\text{social component}} \tag{3}$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \tag{4}$$

where w is the inertia weight and it is a decreasing function of the iteration index.

The velocity vector in Equation (3) consists of three terms that determine the next position:

1. Previous velocity: This is the stored velocity from the previous iteration to regulate each particle from making severe changes in its direction between consecutive iterations.
2. The cognitive component: This term represents the attraction force that each particle has toward its best position achieved based on its own flying experience.
3. The social component: This term corresponds to each particle tendency to be attracted toward the best position discovered among the entire individuals in a swarm.

To maintain a good balance between the individuality and sociality, c1 and c_2 are typically set to be equal. If c_1 is set greater than c_2, each particle individual performance will be weighed more in Equation (3) and it is more likely that the algorithm will get trapped in local solutions (i.e. the best solution achieved by that individual particle). On the contrary, if c_1 is set less than c_2, that algorithm might fail to converge. The inertia weight parameter introduced in Equation (3) allows the velocity vector to start with larger values, and then it decreases as the iteration index increases to limit any big particle movements towards the end of the optimization process. This modification improves the convergence characteristics significantly. Factors affecting the flying experience of each particle in its search for optimal solution are shown in Fig. 2.

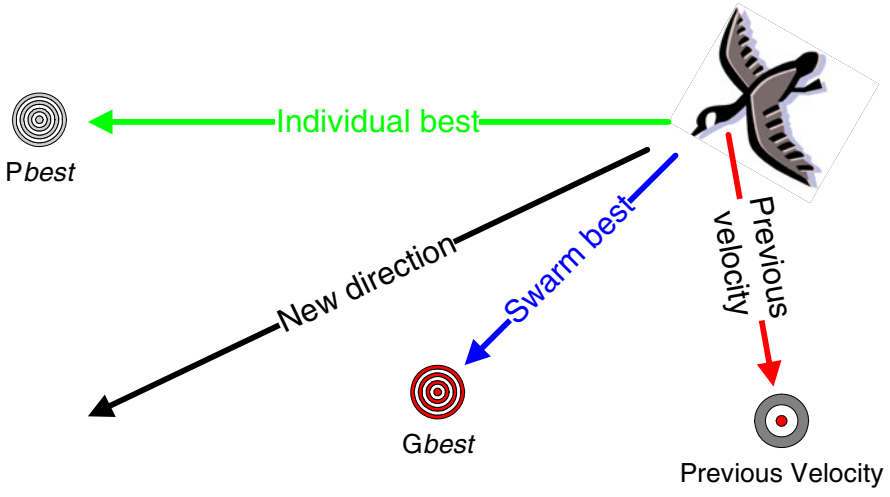


Fig. 2. Influential elements on the particle's movement during its search for an optimum

4 PSO versus Other Optimization Techniques

PSO is a population-based evolutionary technique that has many key advantages over other optimization techniques, for example:

- Unlike deterministic methods, PSO is a non-gradient, derivative-free method which gives the PSO the flexibility to deal with objective functions that are not necessarily continuous, convex or differentiable,
- PSO does not use derivative information (1st and/or 2nd order) in its search for an optimal solution, instead it utilizes the fitness function value to guide the search for optimality in the problem space,
- PSO, by utilizing the fitness function value, eliminates the approximations and assumption operations that are often performed by the conventional optimization methods upon the problem objective and constraint functions,
- Due to the stochastic nature of the PSO method, PSO can be efficient in handling special kinds of optimization problems which have an objective function that has stochastic and noisy nature,
- The quality of a PSO obtained solution, unlike deterministic techniques, does not depend on the initial solution,
- The PSO is a population-based search method that enables the algorithm to evaluate several solutions in a single iteration which, in turn, minimizes the likelihood of the PSO getting trapped in local minima,
- The PSO algorithm is flexible enough to allow hybridization and integration with any other method if needed, whether deterministic or heuristic,

- Unlike many other metaheuristic techniques, PSO has fewer parameters to tune and adjust,
- Overall the PSO algorithm is simple to comprehend, and easy to implement and to program since it utilizes simple mathematical and Boolean logic operations.

On the other hand, PSO has some disadvantages that can be summarized as follows:

- There is no solid mathematical foundation for the PSO metaheuristic method,
- It is a highly problem-dependent solution method, as most metaheuristic methods are; for every system the PSO parameters have to be tuned and adjusted to ensure a good quality solution,
- Other metaheuristic optimization techniques have been commercialized through code packages like MATLAB® GADS® Toolbox for GA, GeaTbx® for both GA and Evolutionary Algorithm (EA) and Excel Premium Solver for EP; however PSO- to the knowledge of the authors- has not commercialized yet.
- Compared to GA, EP algorithms, PSO has fewer published books and articles.

A pseudo-code of general PSO algorithm is shown in Fig. 3.

```
Set the algorithm parameters;
For each particle
        Randomly initialize the position vector;
        Randomly initialize the velocity vector;
End
Measure the fitness of each particle;
Store pbest
Store gbest
While the stopping criteria is not met
        For each particle
                Update the velocity and position vectors
                Measure the fitness of the new position vector
                If the new fitness value is better than the previously stored one
                        Store the new position vector as pbest
                        Store the new fitness value
                End
        End
        Determine the particle with lowest fitness value in the search history
        and store its position vector as gbest
End
```

Fig. 3. A pseudo-code of PSO algorithm

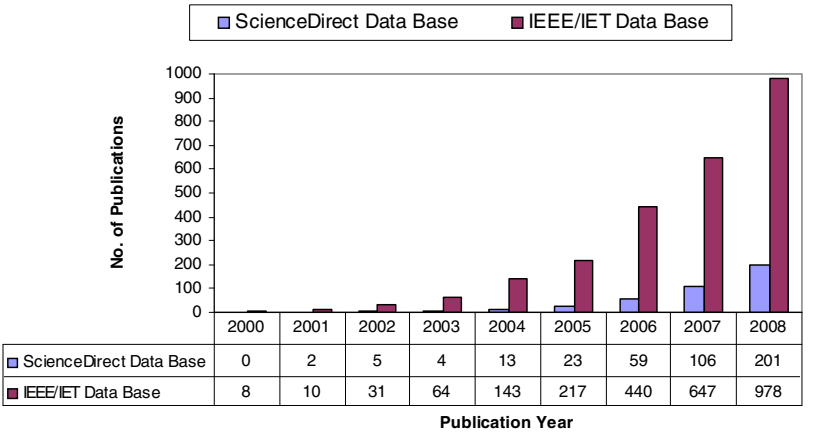
In addition to traditional gradient-based optimization algorithms, there are many other heuristic techniques that compete with PSO such as genetic algorithms, simulated annealing, evolutionary programming, and most recently ant colony optimization. In general, most of these techniques can be used to solve various optimization problems in a similar way to the case of PSO. However, such competing techniques tend to have major drawbacks such as:

- More parameter tuning is required.
- They tend to require more computational time in most cases.
- Heavily involved programming skills are required to develop and modify competing algorithms to suit different classes of optimization problems.
- Some techniques require binary conversion instead of working with direct real valued variables.
- Most of them require a considerable number of population members that would translate to more fitness evaluations.

On the other hand, some advantages of the aforementioned algorithms over PSO are:

- The availability of commercial versions of some algorithms like MATLAB® (genetic algorithm and simulated annealing) and Excel premium solver (evolutionary programming).
- The extensive collection of books and research literature, especially in the case of genetic algorithm and evolutionary programming, that provide broad coverage of these competing methods.

Other heuristic techniques that belong to the same category are summarized in [8]. These techniques have been gaining more popularity mainly because of their robustness, simplicity, and their ability to deal with more exact models instead of making intolerable approximations. The major drawbacks of PSO are the lack of solid mathematical background and failure to theoretically assure global optimal solutions, just like in the case of other metaheuristic optimizers. PSO has been proven to perform well in many standard benchmark optimization problems used by researchers to validate new global optimization techniques [9-12]. Reference [10] is an excellent reference that analyzed and studied the PSO promising convergence characteristics. In [10], Clerc and Kennedy successfully established some mathematical foundations to explain the behavior of a simplified PSO model in its search for an optimal solution. However, further analysis is needed to explain other issues of the PSO like the social influence aspect of the algorithm and generalized rules in how to tune its parameters to suit different optimization problems. In [10], the authors emphasized the need for further future studies by stating "Several kinds of coefficient adjustments are suggested in the present paper, but we have barely scratched the surface and plenty of experiments should be prompted by these findings." Fig. 4 shows the exponentially increasing growth in various research areas with regard to PSO (based on IEEE/IET/ScienceDirect databases).



	2000	2001	2002	2003	2004	2005	2006	2007	2008
ScienceDirect Data Base	0	2	5	4	13	23	59	106	201
IEEE/IET Data Base	8	10	31	64	143	217	440	647	978

Fig. 4. Of publications in IEEE/IET and ScienceDirect Databases since the year 2000

5 Constraint Handling Methods in Evolutionary Algorithms

There are different ways to handle constraints in evolutionary computation optimization algorithms just like in the case of the PSO. The following constraint handling methods are the most commonly used [11]:

1. Preserving feasible solution method: In this method, solutions are initially placed in the feasible search space and remain within this space by adapting an update mechanism that generates only feasible solutions.
2. Infeasible solution rejection method: This approach rejects any solution that violates the feasible search space.
3. Penalty function method: In which a penalty factor is added to the objective once any constraint violation occurs. It transforms the constrained optimization problem to unconstrained one.
4. Solution repair method: This approach converts the infeasible solution to a feasible one by performing special operations.

Selecting the proper constraint handling method is highly reliant on the nature of the problem. Reference [11] indicates that in the solution repair method, the process of reinstating the infeasible solution to a feasible one can be as challenging as solving the original problem. In the penalty function method, the objective function is augmented by adding penalty terms to transform the constrained problem into an unconstrained one. This approach usually encounters a major difficulty in how to properly select penalty factor values. If the penalty factors selected are high, the optimization algorithm will get trapped in local solutions. On the other hand, the algorithm may not be able to detect a feasible solution if the penalty factors are low [13].

6 Survey of Major PSO Applications in Power Systems

The following are the major areas in power systems in which PSO was applied:

6.1 Economic Power Dispatch

El-Gallad *et al.* [14] and Park *et al.* [15] adapted PSO to solve the traditional economic dispatch problem. In both papers, the objective function was formulated as a combination of piecewise quadratic cost functions with non-differential regions instead of using a single convex function for each generating unit. This innovation in problem formulation is due to the incorporation of practical operating conditions like valve-point effects and different fuel types. The system constraints included in reference [14] were system demand balance constraint with network losses incorporated and the generating capacity limits. Park *et al.* did not account for transmission line losses in reference [15] for simplicity. El-Gallad *et al.* added new constraints to the problem formulation in reference [16] by introducing system spinning reserve and generator prohibited operating zones. In this formulation, they included the same constraints as those used in reference [14] and considered a single convex cost function.

In reference [17], a different formulation was proposed by including the generator ramp rate limits in the same problem treated in [16]. In Gaing's work [17], a comparison is made between PSO and genetic algorithm performance in solving the same economic dispatch problem. Gaing introduced a dynamic aspect to the same problem by adding a time-varying system load in addition to accounting for some of the generator operation related restrictions, such as ramping rate limits and prohibited operating zones, while imposing system spinning reserve requirements and line flows as inequality constraints [18]. Victoire and Jeyakumar extended Gaing's research by forming a hybrid optimizer to tackle the same problem [19]. They used SQP to fine-tune PSO search in finding the optimal solution.

Kumar *et al.* included emission aspects of the power dispatching problem [20]. They utilized PSO in solving a multi-objective optimization problem that included both cost and emission functions. They combined the two objective functions by assigning a single price penalty factor to the emission function to form a single objective function. Reference [21] presents improved versions of PSO to solve both convex and non-convex economic dispatch problems that take into account different operational constraints. The main contributions of the proposed approaches are the integration of local random search with PSO and the splitting up of the cognitive term such that both the best and worse particle positions affect the velocity update equation. Wang and Singh formulated a multi-objective emission-economic dispatch problem for a multi-area system [22]. A PSO approach was developed to solve the problem with convex objective functions while accounting for the tie-line transfer limits as additional constraints. Reference [23] presents a hybrid form of PSO and evolutionary programming to solve the economic dispatch while accounting for the valve point loading effects. The hybrid approach showed faster convergence characteristics when compared to the conventional PSO or evolutionary programming.

6.2 Reactive Power Optimization

In this area, PSO was used to optimize the reactive power flow in the power system network to minimize real power system losses. Yoshida *et al.* [24-26] and Fukuyama *et al.* [27] took the initiative of introducing PSO to reactive power optimization. In their problem formulation, the objective was to find the optimal settings of some control variables that would minimize the total real power losses in a network. The control variables are automatic voltage regulator operating values, transformer tap positions, and a number of reactive power compensation equipment subject to equality and inequality constraints. Based on the nature of the control variables, the problem is classified as a mixed-integer nonlinear optimization problem since some variables are continuous while others are discrete. Mantawy and Al-Ghamdi investigated the same problem using a different test system [28].

Miranda and Fonseco appear to be the first to introduce a hybrid PSO approach in this area [29;30]. They combined evolutionary strategies with PSO to improve the robustness of the classical PSO. In [31], Zhao *et al.* combined multi-agent systems with PSO to solve the same problem. Esmin *et al.* considered shunt capacitor banks as the only type of control variables in their problem formulation [32]. They incorporated the tangent vector technique to identify the critical area of power system network where voltage stability might be in danger. Then, they applied PSO to find the "needed" reactive power compensation. A new hybrid method was introduced by Chuanwen and Bompard as they combined PSO with a linear interior point technique to solve a reactive power optimization problem [33]. In their work PSO was used as a global optimizer to search the entire solution space while the linear interior point method acted as a local optimizer to search the space around the optimal solution.

To show the effectiveness of PSO in reactive power control and power losses reduction, it was successfully applied to a practical power system in the province of Heilongjiang in China [34]. This system consists of 151 buses and 220 transmission lines with 71 control variables. A different problem formulation was proposed by Coath *et al.* where they considered reactive power losses minimization as an objective function [35]. They also introduced generator real power outputs as additional control variables. The difference in their problem formulation was mainly due to the inclusion of wind farms as modern integral parts of the power system networks.

6.3 Optimal Power Flow (OPF)

Abido is credited with introducing PSO to solve the OPF problem [36]. In OPF, the goal is to find the optimal settings of the control variables such that the sum of all the generator's cost functions is minimized. The generator real power outputs are considered control variables in addition to the other control variables considered previously in reactive power optimization problems. PSO was effective in dealing with this complex optimization problem that has various equality and inequality constraints and both continuous and discrete variables. In a different approach to the problem, Zhao *et al.* solved the highly constrained OPF optimization

problem by minimizing a non-stationary multi-agent assignment penalty function [37]. In this formulation, PSO was used to solve the highly constrained OPF optimization problem in which the penalty values were dynamically modified in accordance with system constraints. In reference [38], the passive congregation concept was incorporated in PSO to solve the OPF problem. This hybrid technique improved the convergence characteristics over the traditional PSO in solving the same OPF problem. Wang *et al.* developed a modified PSO to solve the OPF problem with the objective being the minimization of the quadratic fuel cost function [39]. The proposed algorithm mainly relied on the idea of randomly exchanging information among the entire swarm rather than only the best member in the swarm. The environmental-economic transaction planning problem in the electricity market was formulated as a multi-objective OPF in reference [40]. A multi-objective PSO algorithm was developed to solve the problem via a non-stationary multi-stage assignment penalty function. Different versions of PSO were developed in reference [41] in an attempt to construct a comparison of their performance with regard to the OPF. The objective functions selected in this study were the real power losses and voltage profile improvement. Gaing introduced an enhanced PSO to solve a multi-objective OPF problem with the objective functions being the fuel cost, real power losses, and voltage deviation [42].

6.4 Controller Design

In references [43] and [44], PSO was employed to find the optimal settings of power system stabilizer parameters. The problem was formulated as one of min-max optimization of two eigenvalue-based objective functions. Okada *et al.* went along the same lines when they used PSO to optimally design a fixed-structure controller to enhance the stability of power systems [45]. In this work, the authors' goal was to find the global optimal solution of a multimodal optimization problem. PSO was also used in optimizing the feedback controller gains. Al-Musabi *et al.* made use of PSO in finding optimal controller gain values for a load frequency problem of a single area power system [46]. Abdel-Magid and Abido extended PSO usage in this area when they enlarged the control system to two areas [47]. In their work, they considered two types of controllers namely an integral controller and a proportional plus integral controller. Juang and Lu combined the genetic algorithm with PSO in reference [48] to perform the same optimization process as in [47] on a fuzzy proportional-integral-controller. Ghoshal augmented the problem by trying to find the optimal proportional-integral-derivative controller gains of a three area power system [49]. He tackled the problem using PSO in addition to other heuristic techniques. Lu and Juang applied PSO to design a fuzzy controller for a thyristor-controlled series capacitor to enhance the transient stability of flexible alternating current transmission systems (FACTS) [50].

6.5 Neural Network Training

Neural Networks emerged as a valuable artificial intelligence tool in many areas in electric power systems. El-Gallad *et al.* used PSO to train a Neural Network for

power transformer protection [51]. The objective was to develop a model that would be able to intelligently distinguish between magnetizing inrush current and internal fault current in power transformers. PSO was employed to improve the accuracy and the execution time of the identification process. Hirata *et al.* used PSO to determine the optimal connection weights of a neural network model used to improve stability control of power systems [52]. They formulated the optimization problem as a min-max problem with an objective function that has non-differential and discontinuous nature. Kassabalidis *et al.* integrated PSO with a Neural Network to identify the dynamic security border of power systems under a deregulated power system environment [53].

6.6 Other Electric Power System Areas

In [54] and [55], the performance of PSO was explored in the area of electric power quality by improving the process of feeder reconfiguration. The problem was formulated as a nonlinear optimization problem with non-differentiable characteristics. Victoire and Jeyakumar combined PSO, sequential-quadratic-programming, and tabu-search to form a hybrid technique to solve the unit commitment combinatorial optimization problem [56]. In the area of short-term load forecasting, Huang *et al.* were able to identify the autoregressive moving with the exogenous variable model using PSO [57]. Slochanal *et al.* and Kannan *et al.* introduced PSO in the area of generation expansion planning in references [58] and [59] to solve discrete nonlinear optimization problems. They used it in [58] to maximize the profit of a generating utility subject to certain market conditions and various system constraints. In [59], PSO was employed to minimize the capital and operation cost of the generation expansion planning problem. Also in this area, PSO was utilized in solving the expansion planning problem of a transmission line network [60].

Koay and Srinivasan solved the multi-objective generator maintenance scheduling problem by creating a hybrid technique by means of combining PSO with evolutionary strategies in reference [61]. In power system reliability studies, PSO was applied to feeder-switch relocation problems in a radial distribution system [62]. The authors in reference [62] used PSO to allocate the most appropriate positions to place sectionalized devices in distribution lines. The objective function of this problem is categorized as nonlinear with non-differentiable characteristics. In reference [63], applications of PSO in finding optimal operation settings of a system composed of distributed generators and energy storage systems were illustrated. Naka *et al.* and Fukuyama formed hybrid techniques by combining PSO with other heuristic techniques to improve the performance of a distribution of state estimator in [64] and [65] respectively. PSO was later applied to solve short term hydroelectric system scheduling problems in reference [66]. The problems in references [64-66] are formulated as continuous nonlinear optimization problems. Yu *et al.* applied PSO to tackle the discrete optimal capacitor placement problem in a noisy environment [67].

7 Application of PSO in Optimal Operation of Power Systems OPF Example

Power engineers require special tools to optimally analyze, monitor, and control different aspects of power systems operation and planning. Most of these tools are properly formulated as some sort of optimization problems. The optimal power flow (OPF) is the backbone tool that has been extensively researched since its first introduction in the early 1960's [68;69]. It appears that the commonly known term "optimal power flow" was first introduced by Dommel and Tinney in 1968 [70].

Originally, the OPF was formulated as a natural extension of the traditional economic dispatch. Differences between the two optimization functions exist even though both of them may share the same objective function. In economic dispatch, the entire power network is reduced to a single equality constraint. By contrast, all major elements of the modeled system are explicitly presented in the OPF problem. The generic term "OPF" is no longer associated exclusively with the extended economic dispatch calculation. Rather, it presents a wide range of optimization problems commonly formulated in power systems related studies. OPF studies are evolving over time from its basic form to cope with the continuous changes that are taking place in power systems. Deregulation of the electric power industry, advances being made in the area of power electronics, and the environmental regulations that are being imposed on power plants are some of the main factors that have played major role in constantly reformulating the OPF. The historical development of the OPF is closely correlated with the advances made in the area of numerical optimization techniques [71]. Researchers have attempted to apply most optimization techniques to solve the OPF.

The purpose of OPF is to find the optimal settings of a given power system network that optimize a certain objective function while satisfying its power flow equations, system security, and equipment operating limits. Different control variables are manipulated to achieve an optimal network setting based on the problem formulation. The main control variables typically used in optimizing the OPF are as follows:

- Generators' real power outputs and voltages.
- Transformer tap changing settings.
- Phase shifters settings and placement for expansion planning.
- Switched capacitors and reactors.
- FACTS devices settings and placement for expansion planning.

A main obstacle of the OPF problem is the nature of the control variables since some of them are continuous (e.g. real power outputs and voltages) and others are discrete (e.g. transformer tap setting, phase shifters, and reactive injections). The presence of discrete variables makes the optimization problem a non-convex one, which in turn complicates the solution methodology. Objective functions of the OPF studies can be arranged into two main categories: traditional and newly emerged objectives. The latter group has emerged mainly as a result of restructuring the electric power industry. Fig. 5 summaries most used objectives in the OPF

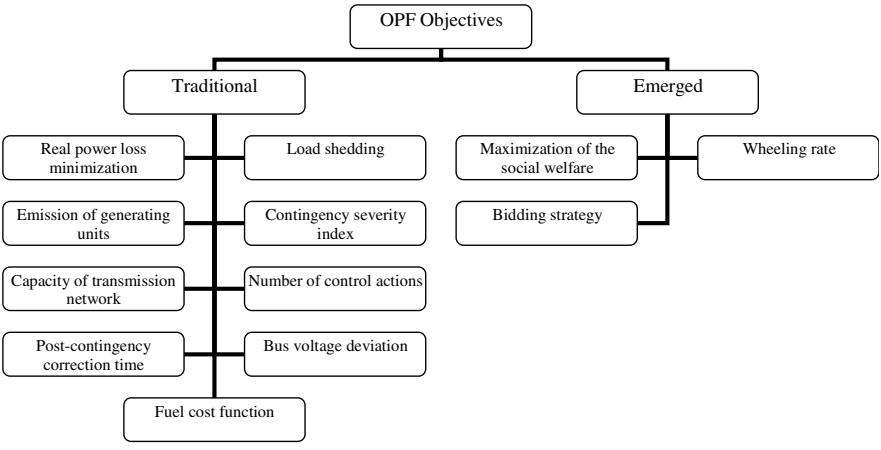


Fig. 5. Most commonly used objective functions in the OPF studies

studies. It is noted from this study that the most commonly used objective is the minimization of the overall fuel cost function (convex and non-convex).

Researchers proposed different mathematical formulations of the OPF problem, which can be broadly classified as follows:

1. Linear programming problem in which objectives and constraints are given in linear forms with continuous control variables.
2. Non-linear programming problem where either objectives or constraints or both combined are non-linear with continuous control variables.
3. Mixed integer linear and non-linear programming problems when control variables are both discrete and continuous.

Many conventional optimization techniques were developed to solve the OPF problem, the most popular being linear programming, sequential quadratic programming, generalized reduced gradient method, and the Newton method. References [72-74] offer a complete list of the most commonly used conventional optimization algorithms with regard to the OPF. Despite the fact that some of these techniques have excellent convergence characteristics and various among them are widely used in the industry, some of their drawbacks are:

1. Convergence to the global or local solution is highly dependant on the selected initial guess, i.e. they might converge to local solutions instead of global ones if the initial guess happens to be in the vicinity of a local solution.
2. Each technique is tailored to suit a specific OPF optimization problem based on the mathematical nature of the objectives and/or constraints.
3. They are developed with some theoretical assumptions, such as convexity, differentiability, and continuity, among other things, which may not be suitable for the actual OPF conditions.

The rapid developments of recent computational intelligence tools have attracted many researchers to employ them in solving the ever changing OPF. Reference [75]

provides an extensive coverage of the major research work conducted in applying modern computational tools to the OPF problem.

8 Problem Formulation

The desired goal of the OPF is to optimize a certain objective subject to different sets of equality and inequality constraints. The problem can be formulated as follows:

$$Min \quad F(x,u) \tag{5}$$

Subject to

$$g(x,u) = 0 \tag{6}$$

$$h_{\min} \le h(x,u) \le h_{\max} \tag{7}$$

where vector x denotes the dependant or state variables of a power system network that contains the slack bus real power output (P_{G1}), voltage magnitudes and phase angles of the load buses (V_{Lk},θ_{Lk}), and generator reactive power outputs (Q_G). Vector u represents both integer and continuous control variables that consist of real power generation levels (P_{GN}) and voltage magnitudes ($|V_{GN}|$), transformer tap setting (T_k), and reactive power injections (Q_{Ck}) due to volt-amperes reactive (VAR) compensations; i.e.

$$u = \begin{bmatrix} \overbrace{u_{continuous}, u_{discrete}}^{OPF\ \ Control\ \ Variables} \end{bmatrix} \tag{8}$$

where

$$u_{continuous} = \left[P_{G2} \ldots P_{GN}, V_{G2} \ldots V_{GN}\right] \text{ and } u_{discrete} = \left[T_1 \ldots T_N, Q_{C1} \ldots Q_{CN}\right] \tag{9}$$

In this example, minimization of different objectives is considered to examine the performance of the proposed algorithm. The objective functions taken into considerations are fuel emission, fuel cost, and the network real power losses. Each objective is briefly described as follows:

8.1 Fuel Emission

Fossil based thermal plants are considered a major player in the pollution crisis that we are facing nowadays. The industrial growth led to greater demands to generate more electricity. Consequently, the emission of these generating units is gradually building up in the atmosphere which is having a severe impact on our environment. One way to cope with this problem is to dispatch electric power with emission considerations. The objective of fuel emission dispatch problem is to minimize the total emission of all thermal units by allocating optimal control settings while satisfying various network operation constraints. Fuel emission of a number of generating units can be modeled mathematically as follows:

$$F_1 = \sum_{i=1}^{N} (\alpha_i + \beta_i P_i + \gamma_i P_i^2) \text{ ton/hr} \tag{10}$$

8.2 Fuel Cost

The aim of the fuel cost dispatch problem is to allocate the best network settings that minimize the overall fuel cost function while imposing all network constraints. Conventionally, the overall fuel cost function for a number of thermal generating units can be modeled by a quadratic function (convex and differentiable) as follows:

$$F_2 = \sum_{i=1}^{i=N} (a_i + b_i P_i + c_i P_i^2) \text{ \$/hr} \tag{11}$$

However, this model ignores the valve point loading that introduces rippling effects to the actual input-output curve. Equation (11) is modified by adding an additional sine term to account for the valve effects in this manner [76]:

$$F_2 = \sum_{i=1}^{i=N} \left[a_i + b_i P_i + c_i P_i^2 + \left| e_i \sin(f_i (P_i^{\min} - P_i)) \right| \right] \text{ \$/hr} \tag{12}$$

This more accurate modeling adds more challenges to most derivative-based optimization algorithms in finding the global solution since the objective is no longer convex nor differentiable every where. Fig. 6 shows the shape of the fuel cost function with the valve loading effects included.

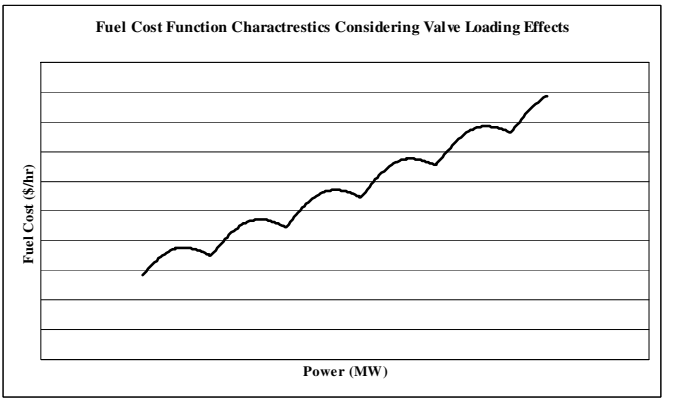


Fig. 6. The generator input-output curve considering the valve point effects

8.3 Real Power Losses

With this objective, all control settings are adjusted such that the total real power losses are minimized. Power losses can be modeled as follows:

$$P_L = \sum_{k=1}^{N_L} g_k \left[\left| V_i \right|^2 + \left| V_j \right|^2 - 2\left| V_i \right| \left| V_j \right| \cos(\delta_i - \delta_j) \right] \tag{13}$$

where N_L is the number of transmission lines in the system, g_k is the conductance of the line k connecting buses i and j, and the bus voltage is represented in polar form by $|V|$ and δ.

The OPF problem has two categories of constraints:

1. Equality Constraints:
These are the sets of nonlinear power flow equations that govern the power system, i.e.

$$P_{G_i} - P_{D_i} - P_{T_i}(\mathbf{V}, \boldsymbol{\delta}) = 0 \tag{14}$$

$$Q_{G_i} - Q_{D_i} - Q_{T_i}(\mathbf{V}, \boldsymbol{\delta}) = 0 \tag{15}$$

where P_{Gi} and Q_{Gi} are the generated real and reactive power at bus i respectively, the load demand at the same bus is represented by P_{Di} and Q_{Di}, and subscript T denotes the total sending and receiving power at each bus.

2. Inequality Constraints:
These are the set of continuous and discrete constraints that represent the system operational and security limits like the bounds on:

1. The generators real and reactive power outputs;

$$P_{G_i}^{\min} \le P_{G_i} \le P_{G_i}^{\max}, \ i = 1, \ldots, G_N \tag{16}$$

$$Q_{G_i}^{\min} \le Q_{G_i} \le Q_{G_i}^{\max}, \ i = 1, \ldots, G_N \tag{17}$$

2. Voltage magnitudes at each bus in the network;

$$V_i^{\min} \le V_i \le V_i^{\max}, \ i = 1, \ldots, N \tag{18}$$

3. The discrete transformer tap settings;

$$T_i^{\min} \le T_i \le T_i^{\max}, \ i = 1, \ldots, T_N \tag{19}$$

4. The discrete reactive power injections due to capacitor banks;

$$Q_{C_i}^{\min} \le Q_{C_i} \le Q_{C_i}^{\max}, \ i = 1, \ldots, C_N \tag{20}$$

Note that P_{Gi}, Q_{Gi}, and V_i are continuous variables while T_i and Q_{Ci} are discrete ones.

5. The transmission lines loading;

$$S_{L_i} \le S_{L_i}^{\max}, \ i = 1, \ldots, L_N \tag{21}$$

Additional inequality constraints may include prohibited zones of the generating units, measures of transient stability, electromagnetic field levels …etc.

9 Simulation Results and Discussion

PSO algorithm was implemented in MATLAB® computing environment and the standard IEEE 30-bus test system was used to validate its potential. The test system consists of six generating units interconnected with 41 branches of a transmission network to serve a total load of 189.2 MW and 107.2 Mvar as shown in Fig. 7. Detailed description of the system's data is presented in the Appendix [77].

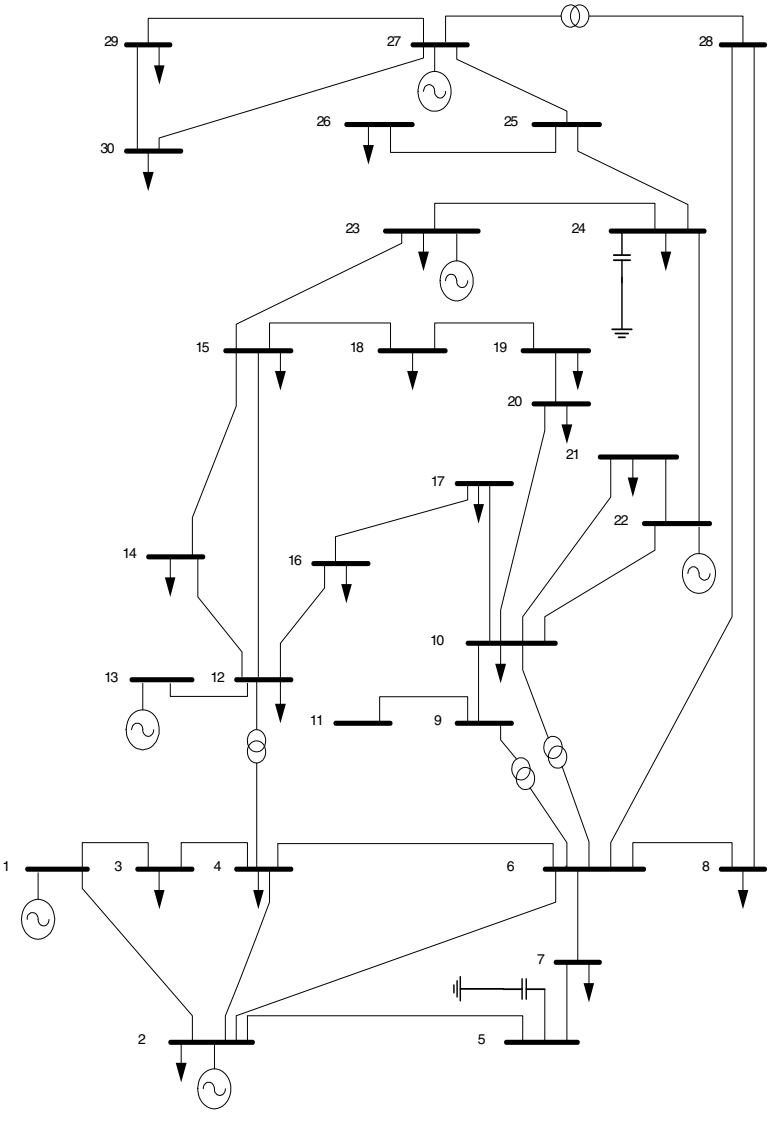


Fig. 7. A single line diagram of IEEE 30-bus standard test system

Table 1. A Study of Tuning PSO Parameters

Parameter		Ave	Min	Max	St. Dev.	Other HPSO Parameters
C_1, C_2	0.10	582.396	577.831	588.311	2.950	No. of Particle = 10
	0.25	578.845	576.501	584.893	2.220	Max. Velocity = 1
	0.50	577.569	575.835	583.552	1.761	Max. Iterations =30
	0.75	576.849	575.841	579.132	0.921	
	1.00	**576.721**	**575.461**	**579.343**	**0.931**	
	1.25	576.971	575.659	578.876	0.932	
	1.50	576.939	575.809	578.321	0.935	
	1.75	577.129	575.526	579.325	0.941	
	2.00	578.266	576.882	581.483	1.201	
	2.50	580.053	576.583	583.388	1.977	
Number of Particles	5	579.679	576.086	586.938	2.910	$C_1 = C_2 = 1.0$
	10	576.721	575.461	579.343	0.931	Max. Velocity = 1
	20	**575.872**	**575.392**	**577.514**	**0.521**	Max. Iterations =30
	30	575.792	575.392	576.788	0.351	
Maximum Velocity	0.01	575.9490312	575.4180022	576.5178288	0.355682775	$C_1 = C_2 = 1.0$
	0.1	**575.7038747**	**575.4107523**	**576.3385352**	**0.262708898**	No. of Particle = 20
	0.25	575.8114492	575.4135668	576.677822	0.342970307	Max. Iterations =30
	0.5	576.3442547	575.4562268	578.8141952	1.064837346	
	1	575.8724351	575.3922643	577.5141537	0.520501034	
	2	576.0249806	575.423816	577.553081	0.563978232	

Note that the original system has two capacitors banks installed at bus 5 and 24 with ratings of 19 and 4 Mvar respectively. A series of experiments were conducted to properly tune the PSO parameters to suit the targeted OPF problem. Considering the quadratic fuel cost function as an objective, Table 1 shows the PSO outcomes as a result of varying its parameters. To quantify the results, 50 independent runs were executed for each parameter variation. The most noticeable observation from this groundwork is that the optimal settings for c_1 and c_2 are found to be 1.0. These values are relatively small since most of the values reported in the literature are in the range of 1.4-2. The best settings for number of particles and particle's maximum velocity are 20 and 0.1 respectively. Note that increasing number of particles more than 20 will improve the solution accuracy slightly at the expense of increasing the computation time significantly.

Inertia weight is kept fixed throughout the simulation process between upper and lower bounds of 0.9 and 0.4 respectively since changing its values did not have great impact in improving the convergence characteristics. The same parameters were suitable to cases 1 and 2 below. Once the PSO best parameters are set, the following cases were considered to test the proposed approach:

Case 1: The quadratic emission and fuel cost functions in equations (10) and (11) were minimized considering only the continuous control variables, i.e. real power outputs and voltages at voltage-controlled buses. A comparison of results obtained using the PSO to those obtained using MATPOWER®, MATLAB®-based software that uses sequential quadratic programming (SQP) to solve the OPF, are shown in Table 2. MATPOWER is capable of solving the OPF when the objective is

Table 2. Comparison between PSO and SQP for CASE 1

	Fuel Cost ($/hr)		Emission (ton/hr)	
Method	SQP	PSO	SQP	PSO
P_{g1}	41.54	43.611	24.88	24.016
P_{g2}	55.4	58.060	28.82	27.601
P_{g13}	16.2	17.555	33.05	30.181
P_{g22}	22.74	22.998	33.06	34.441
P_{g23}	16.27	17.056	26.25	30.000
P_{g27}	39.91	32.567	45.27	45.202
V_1	0.982	1.000	1.033	1.000
V_2	0.979	1.000	1.03	1.001
V_{13}	1.064	1.059	1.1	1.064
V_{22}	1.016	1.012	1.023	1.023
V_{23}	1.026	1.021	1.054	1.043
V_{27}	1.069	1.037	1.068	1.048
Objective	576.892	575.411	284.966	282.628
P_{losses}	2.860	2.647	2.130	2.240

represented in polynomial form and is only capable of handling continuous variables. Results clearly indicate that PSO achieved better solution in both cases.

Case 2: The test system is modified by introducing 4 tap-changing transformers between buses 6-9, 6-10, 4-12, and 27-28. The operating range of all transformers is set between 0.9-1.05 with a discrete step size of 0.01. Also, the capacitor banks at buses 5 and 24 are also considered as new discrete control variables with a range of 0-40 Mvar and a step size of 1. With this modification, the problem now has both continuous and discrete control variables that can be troublesome to most conventional optimization methods. In addition to the objectives considered in case 1, the total real power losses is also introduced as new objective in this case. Table 3 summarizes the results of each minimization problem along with the best solution vector achieved.

Case 3: Since PSO is capable of handling optimization problems in which the objective is not required to be convex or differentiable, the fuel cost function is augmented with an additional sine term as in (12). This addition increases the degree of non-smoothness of the objective function significantly as depicted in Fig. 8. It shows the shape of the fuel cost function of two generation units, which represents part of the overall problem, when the valve point effects are accounted for and in the absence of any constraints. Note that even when considering only two units, the shape of the objective is highly non-convex with multiple peaks and non-differentiable valleys. Note that the fuel cost coefficients are modified to create more challenging objective function shape within the permissible operating

Table 3. Results of Different Objective Minimization When both Discrete and Continuous Variables are Considered (CASE 2)

	Fuel Cost ($/hr)	Emission (ton/hr)	P_{losses} (MW)
P_{g1}	42.180	24.032	7.057
P_{g2}	57.013	27.333	50.131
P_{g13}	17.305	29.817	39.888
P_{g22}	22.025	33.895	45.575
P_{g23}	17.872	29.993	19.116
P_{g27}	35.060	46.202	28.963
V_1	1.000	1.000	1.000
V_2	0.999	1.002	0.950
V_{13}	1.061	1.098	1.100
V_{22}	1.071	1.041	1.091
V_{23}	1.076	1.073	1.093
V_{27}	1.100	1.084	1.093
Q_{C5}	4.000	2.000	9.000
Q_{C24}	8.000	9.000	9.000
$T_{6\text{-}9}$	0.900	0.970	0.900
$T_{6\text{-}10}$	0.950	0.930	0.950
$T_{4\text{-}12}$	0.930	1.020	0.920
$T_{27\text{-}28}$	0.950	0.990	0.980
Objective	574.143	282.218	1.540

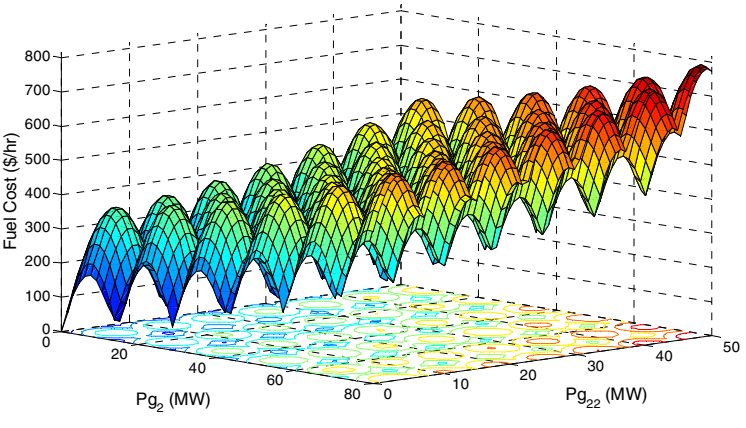


Fig. 8. The high degree of non-convexity in the shape of the objective once valve loading effects are included

Table 4. Results of Case 3 under Different Swarm's Size

	20	30	100
P_{g1}	47.068	47.095	47.126
P_{g2}	42.911	42.359	71.366
P_{g13}	8.790	35.902	8.972
P_{g22}	44.728	37.359	37.391
P_{g23}	8.983	8.826	8.993
P_{g27}	42.044	20.959	20.777
V_1	1.000	1.000	1.000
V_2	1.099	1.009	1.097
V_{13}	1.091	1.017	1.037
V_{22}	1.087	1.082	0.982
V_{23}	1.048	1.057	1.048
V_{27}	1.029	1.080	1.088
Q_{C5}	33.000	16.000	29.000
Q_{C24}	35.000	15.000	12.000
$T_{6\text{-}9}$	1.040	1.010	1.020
$T_{6\text{-}10}$	1.010	1.000	0.990
$T_{4\text{-}12}$	1.040	0.990	1.020
$T_{27\text{-}28}$	0.990	1.030	1.040
Objective	658.416	645.333	615.250

range. In this case, more number of particles is needed to explore this complex solution hyperspace efficiently. Table 4 tabulated the results obtained using different swarm's size. Increasing the swarm's size improved the PSO performance in achieving better results at the expense of computational time.

To demonstrate the consistency and robustness of the developed algorithm, 20 independent runs were conducted for each case to measure the frequency of reaching the optimal or near optimal solution while maintaining the same stopping criteria (maximum iterations of 30). Results and computation time are shown in Table 5 that reflect the steady performance of PSO in solving the OPF problem.

It is evident that in case 1, even the PSO worse performance outperformed SQP in both fuel cost and emission minimization. However, it is noted that in case 3, there was a noticeable deviation between best and worse run when incorporating the valve loading effects into the fuel cost function. This is due to the highly non-smooth feasible region as a result of adding sine terms to the quadratic functions. Similar deviations were noted in earlier work conducted in [78;79] when considering the valve loading effects.

Table 5. Statistical Data for All Testing Cases

		Mean	Best	Worse	Standard Deviation	Average Time (s)/iter
Case 1	Fuel Cost ($/hr)	575.704	575.411	576.339	0.263	5.974
	Emission (ton/hr)	283.110	282.628	283.874	0.386	6.172
Case 2	Fuel Cost ($/hr)	575.228	574.143	576.485	0.550	9.564
	Emission (ton/hr)	283.072	282.218	284.179	0.565	6.781
	P_{losses} (MW)	1.688	1.540	2.018	0.123	10.116
Case 3	20 Particles	744.306	658.416	849.511	61.224	8.519
	30 Particles	734.342	645.333	897.349	70.800	6.712
	100 Particles	677.222	615.250	753.868	42.461	22.870

10 Summary

This chapter covers the basics behind PSO theory and recent developments that have been made to enhance its overall performance. Differences between PSO and other optimization techniques as well as summary of PSO applications in power systems are presented. It highlights many applications in which PSO was successfully applied, yet it reveals some additional unexplored areas where it can be further employed like protection, restoration, etc. Also, deregulating all major parts of the electric power industry led to new operation philosophy to emerge that will reformulate many optimization problems. This will justify using the PSO to tackle such problems. An OPF example is used to illustrate PSO searching capabilities and to demonstrate its potential as a new competitor to other optimization techniques. Future PSO development is anticipated to mainly focus on the theoretical investigations of the global convergence characteristics, hybridization with other optimization techniques, multi-objective optimization, and employment in new applications. PSO is still in its infancy and further development and research are needed to enhance its overall performance characteristics.

Appendix

Table A1. Characteristics of the Generating Units

Generator	1	2	3	4	5	6
a	0	0	0	0	0	0
b	2	1.75	1	3.25	3	3
c	0.02	0.0175	0.0625	0.00834	0.025	0.025
e	300	200	150	100	200	200
f	0.2	0.22	0.42	0.3	0.35	0.35
α	0.04091	0.02543	0.04258	0.05326	0.04258	0.06131
β	-0.05554	-0.06047	-0.05094	-0.03550	-0.05094	-0.05555
γ	0.06490	0.05638	0.04586	0.03380	0.04586	0.05151
Pmin(MW)	0	0	0	0	0	0
Pmax(MW)	80	80	50	55	30	40
Qmin(Mvar)	-20	-20	-15	-15	-10	-15
Qmax(Mvar)	150	60	62.5	48.7	40	44.7
Bus Number	1	2	22	27	23	13

Table A2. Bus Data for IEEE 30-Bus System

Bus No.	Pd	Qd	Qc	V_{min}	V_{max}	Bus No.	Pd	Qd	Qc	V_{min}	V_{max}
1	0	0	0	0.95	1.1	16	3.5	1.8	0	0.90	1.05
2	22	13	0	0.95	1.1	17	9	5.8	0	0.90	1.05
3	2.4	1.2	0	0.90	1.05	18	3.2	0.9	0	0.90	1.05
4	7.6	1.6	0	0.90	1.05	19	9.5	3.4	0	0.90	1.05
5	0	0	19	0.90	1.05	20	2.2	0.7	0	0.90	1.05
6	0	0	0	0.90	1.05	21	18	11	0	0.90	1.05
7	23	11	0	0.90	1.05	22	0	0	0	0.95	1.1
8	30	30	0	0.90	1.05	23	3.2	1.6	0	0.95	1.1
9	0	0	0	0.90	1.05	24	8.7	6.7	4	0.90	1.05
10	5.8	2	0	0.90	1.05	25	0	0	0	0.90	1.05
11	0	0	0	0.90	1.05	26	3.5	2.3	0	0.90	1.05
12	11	7.5	0	0.90	1.05	27	0	0	0	0.95	1.1
13	0	0	0	0.95	1.1	28	0	0	0	0.90	1.05
14	6.2	1.6	0	0.90	1.05	29	2.4	0.9	0	0.90	1.05
15	8.2	2.5	0	0.90	1.05	30	11	1.9	0	0.90	1.05

Table A3. Branch Data for IEEE 30-Bus System

From	To	R	X	B	From	To	R	X	B
1	2	0	0.1	0	15	18	0.1	0.2	0
1	3	0.1	0.2	0	18	19	0.1	0.1	0
2	4	0.1	0.2	0	19	20	0	0.1	0
3	4	0	0	0	10	20	0.1	0.2	0
2	5	0.1	0.2	0	10	17	0	0.1	0
2	6	0.1	0.2	0	10	21	0	0.1	0
4	6	0	0	0	10	22	0.1	0.2	0
5	7	0.1	0.1	0	21	22	0	0	0
6	7	0	0.1	0	15	23	0.1	0.2	0
6	8	0	0	0	22	24	0.1	0.2	0
6	9	0	0.2	0	23	24	0.1	0.3	0
6	10	0	0.6	0	24	25	0.2	0.3	0
9	11	0	0.2	0	25	26	0.3	0.4	0
9	10	0	0.1	0	25	27	0.1	0.2	0
4	12	0	0.3	0	28	27	0	0.4	0
12	13	0	0.1	0	27	29	0.2	0.4	0
12	14	0.1	0.3	0	27	30	0.3	0.6	0
12	15	0.1	0.1	0	29	30	0.2	0.5	0
12	16	0.1	0.2	0	8	28	0.1	0.2	0
14	15	0.2	0.2	0	6	28	0	0.1	0
16	17	0.1	0.2	0	-	-	-	-	-

References

[1] Kennedy, J., Eberhart, R.: Particle swarm optimization. In: IEEE International Conference on Neural Networks, Perth, Australia, vol. 4, pp. 1942–1948 (1995)

[2] Eberhart, R., Kennedy, J.: A new optimizer using particle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, pp. 39–43 (1995)

[3] Heppner, F., Grenander, U.: A stochastic nonlinear model for coordinated bird flocks. The Ubiquity of Chaos, pp. 233–238 (1990)

[4] Xiaohui, H., Yuhui, S., Eberhart, R.: Recent advances in particle swarm. In: Proceedings of 2004 Congress on Evolutionary Computation, vol. 1, pp. 90–97 (2004)

[5] Eberhart, R.C., Shi, Y.: Guest Editorial Special Issue on Particle Swarm Optimization. IEEE Transactions on Evolutionary Computation 8(3), 201–203 (2004)

[6] Valle, Y., Venayagamoorthy, G.K., Mohagheghi, S., Hernandez, J., Harley, R.G.: Particle Swarm Optimization: Basic Concepts, Variants and Applications in Power Systems. IEEE Transactions on Evolutionary Computation 12(2), 171–195 (2008)

[7] Shi, Y., Eberhart, R.: A modified particle swarm optimizer. In: IEEE World Congress on Computational Intelligence, Alaska, USA, pp. 69–73 (1998)
[8] Song, Y.H., Irving, M.R.: An overview of heuristic optimization techniques for power system expansion planning and design. IEE Power Engineering Journal, 151–160 (2001)
[9] El-Gallad, A.I., El-Hawary, M.E., Sallam, A.A.: Swarming of intelligent particles for solving the nonlinear constrained optimization problem. Engineering Intelligent Systems 9(3), 155–163 (2001)
[10] Clerc, M., Kennedy, J.: The particle swarm - explosion, stability, and convergence in a multidimensional complex space. IEEE Transactions on Evolutionary Computation 6(1), 58–73 (2002)
[11] Coath, G., Halgamuge, S.K.: A comparison of constraint-handling methods for the application of particle swarm optimization to constrained nonlinear optimization problems. In: The 2003 Congress on Evolutionary Computation, Canberra, Australia, vol. 4, pp. 2419–2425 (2003)
[12] Yasuda, K., Ide, A., Iwasaki, N.: Stability analysis of particle swarm optimization. In: The fifth Metaheuristics International Conference, pp. 341–346 (2003)
[13] Zhao, B., Guo, C.X., Cao, Y.J.: Improved particle swarm optimization algorithm for OPF problems. In: IEEE/PES Power Systems Conference and Exposition, New York, USA, pp. 233–238 (2004)
[14] El-Gallad, A.I., El-Hawary, M., Sallam, A.A., Kalas, A.: Swarm intelligence for hybrid cost dispatch problem. In: Canadian Conference on Electrical and Computer Engineering, vol. 2, pp. 753–757 (2001)
[15] Park, J.B., Lee, K.S., Shin, J.R., Lee, K.Y.: A particle swarm optimization for economic dispatch with nonsmooth cost functions. IEEE Transactions on Power Systems 20(1), 34–42 (2005)
[16] El-Gallad, A., El-Hawary, M., Sallam, A., Kalas, A.: Particle swarm optimizer for constrained economic dispatch with prohibited operating zones. In: Canadian Conference on Electrical and Computer Engineering, vol. 1, pp. 78–81 (2002)
[17] Gaing, Z.L.: Particle swarm optimization to solving the economic dispatch considering the generator constraints. IEEE Transactions on Power Systems 18(3), 1187–1195 (2003)
[18] Gaing, Z.L.: Constrained dynamic economic dispatch solution using particle swarm optimization. In: IEEE Power Engineering Society General Meeting, pp. 153–158 (2004)
[19] Victoire, T.A.A., Jeyakumar, A.E.: Reserve Constrained Dynamic Dispatch of Units With Valve-Point Effects. IEEE Transactions on Power Systems 20(3), 1273–1282 (2005)
[20] Kumar, A.I.S., Dhanushkodi, K., Kumar, J.J., Paul, C.K.C.: Particle swarm optimization solution to emission and economic dispatch problem. In: Conference on Convergent Technologies for Asia-Pacific Region, vol. 1, pp. 435–439 (2003)
[21] Selvakumar, A.I., Thanushkodi, K.: A New Particle Swarm Optimization Solution to Nonconvex Economic Dispatch Problems. IEEE Transactions on Power Systems 22(1), 42–51 (2007)
[22] Wang, L., Singh, C.: Reserve-Constrained Multiarea Environmental/Economic Dispatch Using Enhanced Particle Swarm Optimization. In: IEEE Systems and Information Engineering Design Symposium, pp. 96–100 (2006)

[23] Sinha, N., Purkayastha, B.: PSO embedded evolutionary programming technique for nonconvex economic load dispatch. In: IEEE/PES Power Systems Conference and Exposition, pp. 66–71 (2004)
[24] Yoshida, H., Kawata, K., Fukuyama, Y., Takayama, S., Nakanishi, Y.: A particle swarm optimization for reactive power and voltage control considering voltage security assessment. IEEE Transactions on Power Systems 15(4), 1232–1239 (2000)
[25] Yoshida, H., Kawata, K., Fukuyama, Y., Takayama, S., Nakanishi, Y.: A particle swarm optimization for reactive power and voltage control considering voltage security assessment. In: IEEE Power Engineering Society Winter Meeting, vol. 2, pp. 498–504 (2001)
[26] Yoshida, H., Fukuyama, Y., Takayama, S., Nakanishi, Y.: A particle swarm optimization for reactive power and voltage control in electric power systems considering voltage security assessment. In: IEEE International Conference on Systems, man, and Cybernetics, vol. 6, pp. 497–502 (1999)
[27] Fukuyama, Y., Yoshida, H.: A particle swarm optimization for reactive power and voltage control in electric power systems. In: Proceedings of the 2001 Congress on Evolutionary Computation, vol. 1, pp. 87–93 (2001)
[28] Mantawy, A.H., Al-Ghamdi, M.S.: A new reactive power optimization algorithm. In: IEEE Power Tech. Conference Proceedings, vol. 4, pp. 6–11 (2003)
[29] Miranda, V., Fonseca, N.: EPSO-evolutionary particle swarm optimization, a new algorithm with applications in power systems. In: IEEE/PES Transmission and Distribution Conference and Exhibition 2002: Asia Pacific., vol. 2, pp. 745–750 (2002)
[30] Miranda, V., Fonseca, N.: EPSO - best-of-two-worlds meta-heuristic applied to power system problems. In: Proceedings of the 2002 Congress on Evolutionary Computation, vol. 2, pp. 1080–1085 (2002)
[31] Zhao, B., Guo, C.X., Cao, Y.J.: A multiagent-based particle swarm optimization approach for optimal reactive power dispatch. IEEE Transactions on Power Systems 20(2), 1070–1078 (2005)
[32] Esmin, A.A.A., Lambert-Torres, G., Zambroni de Souza, A.C.: A hybrid particle swarm optimization applied to loss power minimization. IEEE Transactions on Power Systems 20(2), 859–866 (2005)
[33] Chuanwen, J., Bompard, E.: A hybrid method of chaotic particle swarm optimization and linear interior for reactive power optimisation. Mathematics and Computers in Simulation 68(1), 57–65 (2005)
[34] Zhang, W., Liu, Y.: Reactive power optimization based on PSO in a practical power system. In: IEEE Power Engineering Society General Meeting, pp. 239–243 (2004)
[35] Coath, G., Al-Dabbagh, M., Halgamuge, S.K.: Particle swarm optimisation for reactive power and voltage control with grid-integrated wind farms. In: IEEE Power Engineering Society General Meeting, pp. 303–308 (2004)
[36] Abido, M.A.: Optimal power flow using particle swarm optimization. International Journal of Electrical Power & Energy Systems 24(7), 563–571 (2002)
[37] Zhao, B., Guo, C.X., Cao, Y.J.: Improved particle swam optimization algorithm for OPF problems. In: IEEE/PES Power Systems Conference and Exposition, pp. 233–238 (2004)
[38] He, S., Wen, J.Y., Prempain, E., Wu, Q.H., Fitch, J., Mann, S.: An improved particle swarm optimization for optimal power flow. In: International Conference on Power System Technology, The Pan Pacific, Singapore, vol. 2, pp. 1633–1637 (2004)

[39] Wang, C., Yuan, H., Huang, Z., Zhang, J., Sun, C.: A modified particle swarm optimization algorithm and its application in optimal power flow problem. In: Proceedings of 2005 International Conference on Machine Learning and Cybernetics, Guangzhou, China, vol. 5, pp. 2885–2889 (2005)
[40] Ma, R., Wang, P., Yang, H., Hu, G.: Environmental/Economic Transaction Planning Using Multiobjective Particle Swarm Optimization and Non-Stationary Multi-Stage Assignment Penalty Function. In: 2005 IEEE/PES Transmission and Distribution Conference and Exhibition: Asia and Pacific, Dalian, China, pp. 1–6 (2005)
[41] Vlachogiannis, J.G., Lee, K.Y.: A Comparative Study on Particle Swarm Optimization for Optimal Steady-State Performance of Power Systems. IEEE Transactions on Power Systems 21(4), 1718–1728 (2006)
[42] Gaing, Z.L.: Constrained optimal power flow by mixed-integer particle swarm optimization. In: IEEE Power Engineering Society General Meeting, San Francisco, USA, pp. 243–250 (2005)
[43] Abido, A.A.: Particle swarm optimization for multimachine power system stabilizer design. In: IEEE Power Engineering Society Summer Meeting, vol. 3, pp. 1346–1351 (2001)
[44] Abido, M.A.: Optimal design of power-system stabilizers using particle swarm optimization. IEEE Transactions on Energy Conversion 17(3), 406–413 (2002)
[45] Okada, T., Watanabe, T., Yasuda, K.: Parameter tuning of fixed structure controller for power system stability enhancement. In: IEEE/PES Transmission and Distribution Conference and Exhibition 2002: Asia Pacific, vol. 1, pp. 162–167 (2002)
[46] Al-Musabi, N.A., Al-Hatnouz, Z.M., Al-Duwaish, H.N., Al-Baiyat, S.: Variable structure load frequency controller using particle swarm optimization technique. In: Proceedings of the 2003 10th IEEE International Conference on Electronics, Circuits and Systems, vol. 1, pp. 380–383 (2003)
[47] Abdel-Magid, Y.L., Abido, M.A.: AGC tuning of interconnected reheat thermal systems with particle swarm optimization. In: Proceedings of the 2003 10th IEEE International Conference on Electronics, Circuits and Systems, vol. 1, pp. 376–379 (2003)
[48] Juang, C., Lu, C.: Power system load frequency control by evolutionary fuzzy PI controller. In: IEEE International Conference on Fuzzy Systems, vol. 2, pp. 715–719 (2004)
[49] Ghoshal, S.P.: Optimizations of PID gains by particle swarm optimizations in fuzzy based automatic generation control. Electric Power Systems Research 72(3), 203–212 (2004)
[50] Chun-Feng, L., Chia-Feng, J.: Evolutionary fuzzy control of flexible AC transmission system. IEE Proceedings on Generation, Transmission and Distribution 152(4), 441–448 (2005)
[51] El-Gallad, A.I., El-Hawary, M., Sallam, A.A., Kalas, A.: Swarm-intelligently trained neural network for power transformer protection. In: Canadian Conference on Electrical and Computer Engineering, vol. 1, pp. 265–269 (2001)
[52] Hirata, N., Ishigame, A., Nishigaito, H.: Neuro stabilizing control based on Lyapunov method for power system. In: Proceedings of the 41st SICE Annual Conference, vol. 5, pp. 3169–3171 (2002)
[53] Kassabalidis, I.N., El-Sharkawi, M.A., Marks II, R.J., Moulin, L.S., ves da Silva, A.P.: Dynamic security border identification using enhanced particle swarm optimization. IEEE Transactions on Power Systems 17(3), 723–729 (2002)
[54] Chang, R.F., Lu, C.N.: Feeder reconfiguration for load factor improvement. In: IEEE Power Engineering Society Winter Meeting, vol. 2, pp. 980–984 (2002)

[55] Shen, C.C., Lu, C.N.: Feeder reconfiguration for power quality requirement and feeder service quality matching. In: IEEE/PES Transmission and Distribution Conference and Exhibition 2002: Asia Pacific., vol. 1, pp. 226–231 (2002)
[56] Victoire, T.A.A., Jeyakumar, A.E.: Unit commitment by a tabu-search-based hybrid-optimisation technique. IEE Proceedings on Generation, Transmission and Distribution 152(4), 563–574 (2005)
[57] Huang, C.M., Huang, C.J., Wang, M.L.: A particle swarm optimization to identifying the ARMAX model for short-term load forecasting. IEEE Transactions on Power Systems 20(2), 1126–1133 (2005)
[58] Mary Raja Slochanal, S., Kannan, S., Rengaraj, R.: Generation expansion planning in the competitive environment. In: International Conference on Power System Technology, vol. 2, pp. 1546–1549 (2004)
[59] Kannan, S., Slochanal, S.M.R., Subbaraj, P., Padhy, N.P.: Application of particle swarm optimization technique and its variants to generation expansion planning problem. Electric Power Systems Research 70(3), 203–210 (2004)
[60] Sensarma, P.S., Rahmani, M., Carvalho, A.: A comprehensive method for optimal expansion planning using particle swarm optimization. In: IEEE Power Engineering Society Winter Meeting, vol. 2, pp. 1317–1322 (2002)
[61] Koay, C.A., Srinivasan, D.: Particle swarm optimization-based approach for generator maintenance scheduling. In: Proceedings of the IEEE Swarm Intelligence Symposium, pp. 167–173 (2003)
[62] Kurutach, W., Tuppadung, Y.: Feeder-switch relocation based upon risk analysis of trees-caused interruption and value-based distribution reliability assessment. In: IEEE Region 10 Conference, vol. C, pp. 577–580 (2004)
[63] Koichi Nara, Y.M.: Particle swarm optimization for fault state power supply reliability enhancement. In: Proc. of the Intelligent System Application to Power Systems (ISAP 2001), pp. 143–147 (2001)
[64] Naka, S., Genji, T., Yura, T., Fukuyama, Y.: A hybrid particle swarm optimization for distribution state estimation. IEEE Transactions on Power Systems 18(1), 60–68 (2003)
[65] Fukuyama, Y.: State estimation and optimal setting of voltage regulator in distribution systems. In: IEEE Power Engineering Society Winter Meeting, vol. 2, pp. 930–935 (2001)
[66] Chuanwen, J., Bompard, E.: A self-adaptive chaotic particle swarm algorithm for short term hydroelectric system scheduling in deregulated environment. Energy Conversion and Management 46(17), 2689–2696 (2005)
[67] Yu, X.M., Xiong, X.Y., Wu, Y.W.: A PSO-based approach to optimal capacitor placement with harmonic distortion consideration. Electric Power Systems Research 71(1), 27–33 (2004)
[68] Squires, R.B.: Economic dispatch of generation directly from power system voltages and admittances. AIEE Transactions on Power Apparatus and Systems PAS-79(3), 1235–1245 (1961)
[69] Carpentier, J.: Contribution e l'étude do Dispatching Economique. Bulletin Society Francaise Electriciens, 431–447 (1962)
[70] Dommel, H.W., Tinney, W.F.: Optimal Power Flow Solutions. IEEE Transactions on Power Apparatus and Systems PAS-87(10), 1866–1876 (1968)
[71] Momoh, J.A.: Electric Power System Applications of Optimization. Marcel Dekker, New York (2001)

[72] Momoh, J.A., Adapa, R., El-Hawary, M.E.: A review of selected optimal power flow literature to 1993. I. Nonlinear and quadratic programming approaches. IEEE Transactions on Power Systems 14(1), 96–104 (1999)
[73] Momoh, J.A., El-Hawary, M.E., Adapa, R.: A review of selected optimal power flow literature to 1993. II. Newton, linear programming and interior point methods. IEEE Transactions on Power Systems 14(1), 105–111 (1999)
[74] Huneault, M., Galiana, F.D.: A survey of the optimal power flow literature. IEEE Transactions on Power Systems 6(2), 762–770 (1991)
[75] AlRashidi, M.R., El-Hawary, M.E.: Applications of computational intelligence techniques for solving the revived optimal power flow problem. Electric Power Systems Research 79(4), 694–702 (2009)
[76] Walters, D.C., Sheble, G.B.: Genetic algorithm solution of economic dispatch with valve point loading. IEEE Transactions on Power Systems 8(3), 1325–1332 (1993)
[77] AlRashidi, M.R., El-Hawary, M.E.: Hybrid Particle Swarm Optimization Approach for Solving the Discrete OPF Problem Considering the Valve Loading Effects. IEEE Transactions on Power Systems 22(4), 2030–2038 (2007)
[78] Kulworawanichpong, T., Sujitjorn, S.: Optimal power flow using tabu search. IEEE Power Engineering Review 22(6), 37–55 (2002)
[79] Coelho, L., Mariani, V.C.: Combining of chaotic differential evolution and quadratic programming for economic dispatch optimization with valve-point effect. IEEE Transactions on Power Systems 21(2), 989–996 (2006)

Application of Evolutionary Optimization Techniques for PSS Tuning

S.P. Ghoshal, A. Chatterjee, and V. Mukherjee

Abstract. In this chapter, bacteria foraging optimization (BFO) and chaotic ant swarm optimization (CASO) are individually considered to tune the parameters of both single-input and dual-input power system stabilizers (PSSs). Conventional PSS (CPSS) and the three dual-input IEEE PSSs (PSS2B, PSS3B, and PSS4B) are optimally tuned to obtain the optimal transient performances. A comparative performance study of these four variants of PSSs is also made. It is revealed by applying either BFO or CASO that the transient performance of dual-input PSS is better than single-input PSS. It is, further, explored that among dual-input PSSs, PSS3B offers superior transient performance. A comparison between the results of the BFO and that of genetic algorithm (GA) is conducted in this study. The comparison reveals that BFO is more effective than GA in finding the optimal transient performance. CASO explores the chaotic and self-organization behavior of ants in the foraging process. A novel concept, like craziness, is introduced in the CASO to achieve improved performance of the algorithm. For on-line, off-nominal operating conditions Sugeno fuzzy logic (SFL) based approach is adopted. On real time measurements of system operating conditions, SFL adaptively and very fast yields on-line, off-nominal optimal stabilizer parameters.

Keywords: Ant colony optimization; bacteria foraging optimization; chaotic ant swarm optimization; genetic algorithm; particle swarm optimization; power system stabilizer; Sugeno fuzzy logic; swarm intelligence.

1 Introduction

MOST of the generators are equipped with voltage regulators to automatically control the terminal voltage. It is reported in the literature [1] that the volt-

S.P. Ghoshal
National Institute of Technology, Durgapur, India
E-mail: spghoshalnitdgp@gmail.com

A. Chatterjee
Asansol Engineering College, Asansol, India
E-mail: nirsha_apurba@rediffmail.com

V. Mukherjee
Indian School of Mines, Dhanbad, India
E-mail: vivek_agamani@yahoo.com

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 325 – 366.
springerlink.com

age regulator action had a detrimental impact upon the dynamic stability of the power system. Oscillations of small magnitude and low frequency often persist for long periods of time and in some cases even limit the power transfer capability. Thus, low frequency oscillations or so-called electromechanical oscillations (EMO) [2] usually occur in large power system and some times cause power system to become unstable. Small-signal analysis tools available today help to characterize both the origin and nature of modes. Classification of the modes of EMO, according to their origin and nature, are of the following types:

- Intra plant mode oscillations,
- Local plant mode oscillations,
- Inter area mode oscillations,
- Control mode oscillations,
- Torsional modes between rotating plant.

Depending upon the unit ratings and the reactance connecting them, machines of the same power generation site oscillate against each other in the range of 2.0-3.0 Hz. This oscillation manifests within the generation plant and is referred to *intra plant*. The very beauty of this type of oscillation is that the rest of the system is unaffected. In *local mode*, one generator swings against the rest of the system at 1.0-2.0 Hz. Two coherent groups of generators, swinging against each other at 1 Hz or less cause *inter area* mode of oscillations. The highlighting feature of this complex phenomenon is the involvement of many parts of the system with highly non-linear dynamic behavior. Generators, purely tuned exciters, governors, high voltage direct current (HVDC) converters, static VAR compensator (SVC) controls are mainly associated with *control mode oscillation*. Complex interaction of transformer tap-changing devices with non-linear load dynamics may lead to the birth of voltage oscillations. *Torsional mode oscillation* [3] in the frequency range of 10-46 Hz in the turbine alternator shaft system is noticed. This mode of oscillation plays its role when multi-stage turbine generator unit is connected with the grid system through a series compensated power transmission channel. A mechanical torsional mode of the shaft system interacts with the series capacitor at the natural frequency of the electrical system. When synchronous frequency minus torsional frequency equals network natural frequency, the shaft resonance appears. Glimpses of power system separations with scale blackouts as laid down in [4] may be recalled. Of all the modes described above, local area mode of oscillation will be investigated in this chapter.

Utility houses accept power system stabilizer (PSS) to generate supplementary control signals for the excitation system. PSS provides modulation of the voltage reference of the automatic voltage regulator (AVR) with proper phase and gain compensation circuitry [5]. Power system dynamic performance is improved by damping of system oscillations. This is a very effective method of enhancing the

small-signal stability performance. Small-signal stability analysis and transient stability analysis are complementary techniques. Both of these are essential for studying rotor-angle stability of power system. To investigate the small-signal stability, the usage of non-linear time domain simulation approach possesses the following significant problems:

- Disturbances will not necessarily excite local controller modes which are unstable or poorly damped,
- The time responses are multi-modal,
- There exists usually insufficient information to identify the originality of the modes present in the time responses or to design remedial measures if they are found to be necessary.

Recently, evolutionary programming and intelligent control techniques are being applied to solve many complex optimization problems in engineering applications. With high speed computing tools, these search methods are increasingly being applied in power system planning, design, operation and control problems. The advantage of these methods is that the objective function need not be explicit or differentiable and nonlinearity or non-convexity is not a problem and minimal damping in the closed loop can be obtained.

Most of the techniques, described in the state-of-the-art literature, are centered on angular speed deviation as single input variable to classical phase compensation PSS. Some of these techniques suffer from complexity of algorithm, heavy computational burden and memory storage requirement. Some suffer from robustness because of choice of limited number of control parameters of stabilizer, limited number of optimization functions and necessity of real time on-line fast changing stabilizer parameters. Some techniques like genetic algorithm (GA), simulated annealing (SA) etc suffer from settings of algorithm parameters and give rise to repeated revisiting of the same suboptimal solutions.

In this chapter, GA and two novel evolutionary optimization techniques like chaotic ant swarm optimization (CASO) and bacteria foraging optimization (BFO) are applied to tune the conventional power system stabilizer (CPSS) and three different varieties of dual input IEEE-PSS. The prime emphasis of this chapter is to present a comparison among the conventional phase compensation single input PSS and three varieties of dual input IEEE-PSS. It is also going to present in this chapter that among the dual input IEEE-PSS family (IEEE-PSS2B, IEEE-PSS3B and IEEE-PSS4B) which dual input IEEE-PSS offers the best transient stabilization performance. For on-line tuning of PSS's parameters, very fast acting Takagi Sugeno fuzzy logic (SFL) is adopted.

The organization of this chapter is as follows. In Section 2, single-machine infinite bus (SMIB) model and various PSS models are given. Problem formulation, PSS parameters tuning by evolutionary optimization techniques are given in Section 3 and Section 4, respectively. Section 5 and Section 6 show application of SFL to PSS parameters on-line tuning and input parameters considered,

respectively. The performance studies of various PSSs are elaborated in Section 7. Classical phase compensation based PSS vis-à-vis IEEE dual input PSS is narrated in Section 8. Finally, conclusions are drawn in Section 9.

2 SMIB Model and Various PSS Models

The SMIB system [6], as considered in the present chapter, is illustrated in Fig. 1. The theoretical basis of the PSS representation may be illustrated with the help of block diagram as depicted in Fig. 2 [6, 7]. As the purpose of PSS is to introduce damping torque component, speed deviation is used as a logical signal to control generator excitation. The block diagram of SMIB system with AVR, thyristor high gain exciter, synchronous generator and PSS are shown in Fig. 3. The synchronous generator with AVR, IEEE type ST1A thyristor excitation system and equivalent transmission line reactance are represented by a two-axis, fourth order model. In Fig. 3, ΔT_e input to PSS loop subsystem is assumed to be zero to configure CPSS [6]. In this figure, PSS loop subsystem constitutes Figs. 4-7 to configure CPSS or dual input PSSs [8] as PSS2B, or PSS3B, or PSS4B respectively.

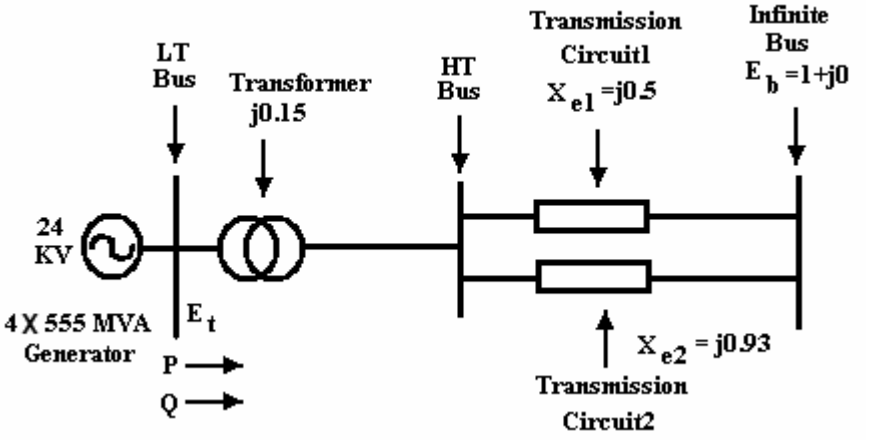


Fig. 1. Single-machine infinite bus test system

CPSS is the single input conventional PSS where $\Delta\omega_r$ is the input and the output is ΔV_{pss}. On the other hand, for dual input PSSs, the two inputs are $\Delta\omega_r$ and ΔT_e and the processed output is ΔV_{pss}. ΔV_{pss} acts as a supplementary control signal to the excitation system. The transfer function of CPSS (Fig. 4) model is given in (1).

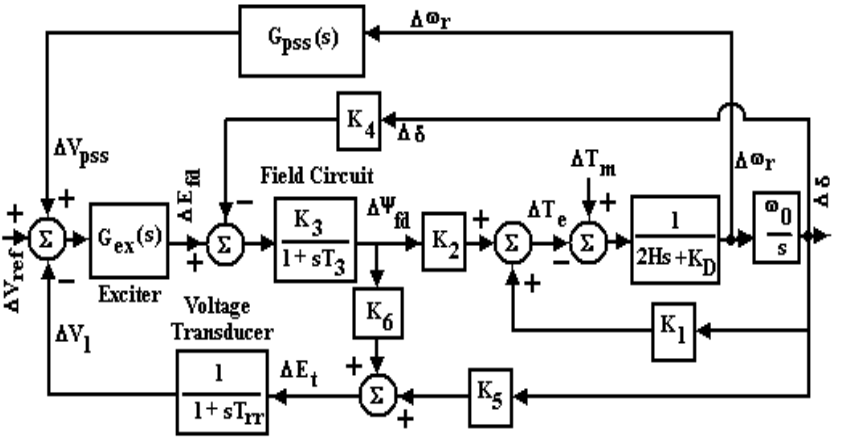


Fig. 2. Block diagram representation with AVR and PSS

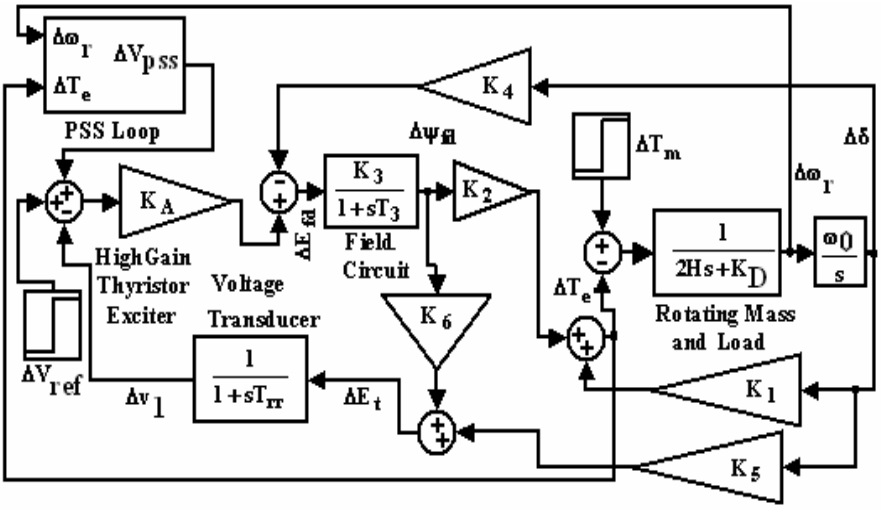


Fig. 3. Block diagram representation of SMIB system with AVR, thyristor high gain exciter, synchronous generator and PSS Loop

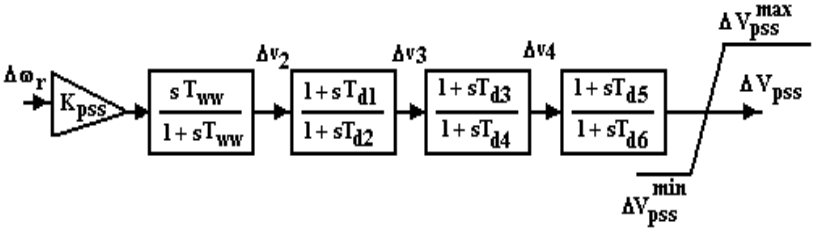


Fig. 4. Block diagram representation of CPSS

$$G_{pss}(s) = \frac{\Delta V_{pss}}{\Delta w_r} = K_{pss}\left(\frac{sT_{ww}}{(1+sT_{ww})}\right)x \left[\frac{(1+sT_{d1})}{(1+sT_{d2})}\frac{(1+sT_{d3})}{(1+sT_{d4})}\frac{(1+sT_{d5})}{(1+sT_{d6})}\right] \quad (1)$$

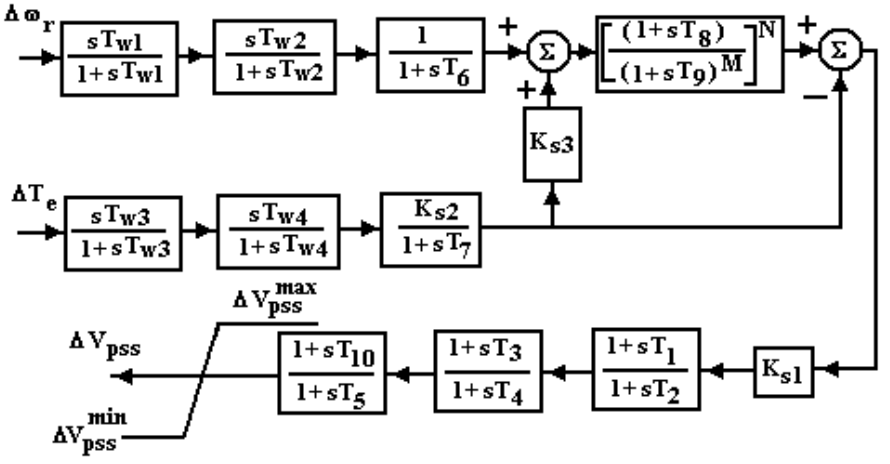


Fig. 5. Block diagram representation of dual input IEEE PSS2B

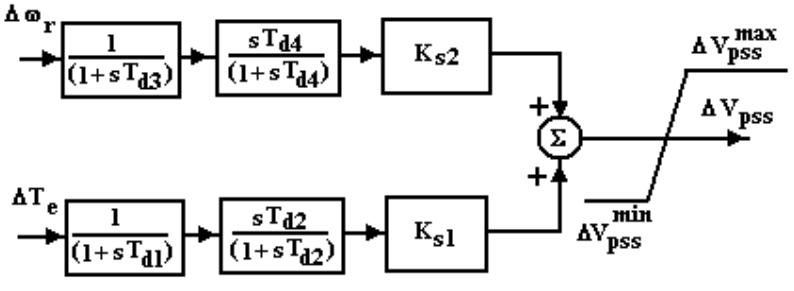


Fig. 6. Block diagram representation of dual input IEEE PSS3B

The second and third term in (1) represents washout term and the lead-lag compensator blocks, respectively. The purpose of using washout block is to allow signals in the range of 0.2-2.0 Hz associated with rotor oscillations to pass unchanged. Lead-lag compensators introduced in the circuit between the exciter input (i.e. PSS output) and the electrical torque provide the phase lead (in some rare cases lag also) for the phase lag. It is important to avoid interaction between the PSS and the torsional modes of vibration. From literature, it is revealed that all modern excitation systems are prone to such interaction as these devices yield

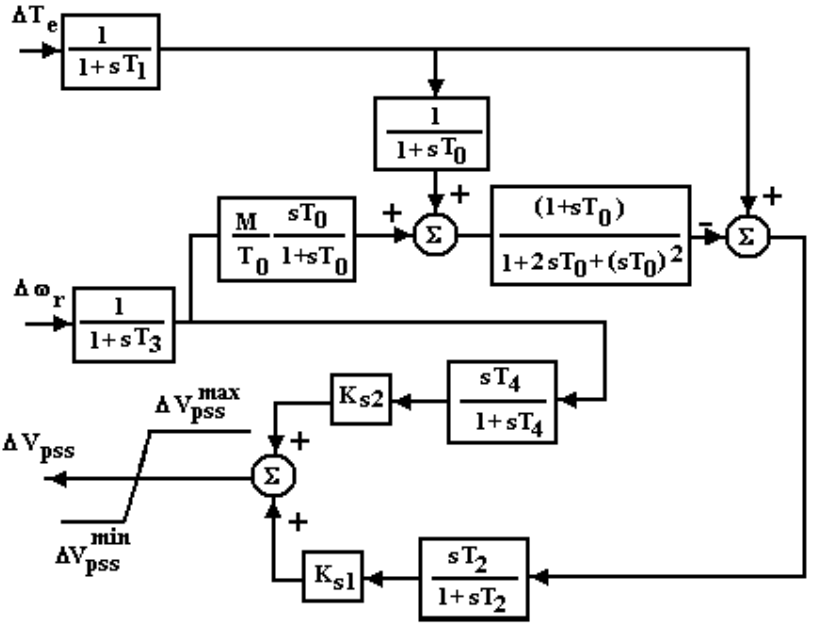


Fig. 7. Block diagram representation of dual input IEEE PSS4B

relatively high gain at high frequencies [9]. Particularly, at light loading conditions of the synchronous generators where the inherent mechanical damping is relatively small, stabilizer-torsional instability may result with a high-response excitation system. Shaft damage does not occur due to such instability but it can cause saturation of the stabilizer output yielding the stabilizer to be ineffective and possibly causing saturation of the voltage regulator, which in turn, results in loss of synchronism and tripping of the unit. Thus, it is imperative that stabilizers do not induce torsional instabilities. A modified form of the model PSS2A published in IEEE standard 421.5 is the computer representation model of PSS2B [8] as illustrated in Fig. 5. To model the stabilizers that incorporate a more complex phase lead function, an additional block with lag time constant T_5 and lead time constant T_{10} can be used. In some specific applications, designers may choose to provide additional low-pass filtering at turbine-generator torsional frequencies, so the model should be structured to permit the use of a zero lead time constant, T_{10}. All computer representation models of the dual input PSS (PSS2B, PSS3B and PSS4B depicted in Figs. 5-7) have two inputs $\Delta\omega_r$ and ΔT_e. Time constants (T_{d2}, T_{d4} in PSS3B model and T_2, T_4 in PSS4B model) represent washout time constants for electrical torque and rotor angular speed respectively. In PSS3B

model the stabilizing signal ΔV_{pss} results from the vector summation of processed signals for electrical torque and angular frequency deviation. By matching the polarity of the gain constants K_{s1} and K_{s2}, the desired amplitude and phase for the stabilizing signal are obtained. The function of the limit values $\Delta V_{pss}^{\max}$ and $\Delta V_{pss}^{\min}$ is to adjust the maximum and minimum allowed, respectively. For PSS4B, the structure of the angular frequency deviation channel, the formulation and limitation of the stabilizing signal are, basically, the same as the structure of the PSS3B model. The conditioning network for accelerating power requires the system constant T_0 (represents the characteristic start up time constant of the synchronous machine) as well as the inertia time constant $M (= 2H)$ for the combined turbine-generator shaft system.

The SMIB system with AVR and exciter along with CPSS, or PSS2B, or PSS3B, or PSS4B is represented by a state matrix A of order eight, seventeen, eight and eleven, respectively. Varying input operating conditions for the present study are active power (P), reactive power (Q), equivalent transmission line reactance (X_e) and Generator's LT side bus voltage (E_t).

3 Problem Formulation

The parameters of the PSS (for CPSS: K_{pss}, T_{d1}, T_{d2}, T_{d3}, T_{d4}, T_{d5}, T_{d6}; for PSS2B: K_{s1}, T_1, T_2, T_3, T_4, T_5; for PSS3B: K_{s1}, K_{s2}, T_{d1}, T_{d2}, T_{d3}, T_{d4}; for PSS4B: K_{s1}, K_{s2}, T_1, T_2, T_3, T_4) are to be so tuned that some degree of relative stability and damping of electromechanical modes of oscillations, minimized undershoot (u_{sh}), minimized overshoot (o_{sh}) and lesser settling time (t_{st}) of transient oscillations of $\Delta\omega_r$ are achieved. So, to satisfy all these requirements, two multi-objective optimization functions, $OF_1()$ and $OF_2()$ which are to be minimized in succession are designed in the following way:

$OF_{11} = \sum_i (\sigma_0 - \sigma_i)^2$ if $\sigma_0 > \sigma_i$, σ_i is the real part of the i$^{\text{th}}$ eigenvalue. The relative stability is determined by $-\sigma_0$. The value of σ_0 is taken as 6.0 for the best relative stability and optimal transient performance.

$OF_{12} = \sum_i (\xi_0 - \xi_i)^2$, if ($\beta_i$, imaginary part of the i$^{\text{th}}$ eigenvalue) > 0.0, ξ_i is the damping ratio of the i$^{\text{th}}$ eigenvalue and $\xi_i < \xi_0$. Minimum damping ratio considered, $\xi_0 = 0.3$. Minimization of this objective function will minimize maximum overshoot.

$OF_{13} = \sum_i \beta_i^2$, if $\sigma_i \geq -\sigma_0$. High value of β_i to the right of vertical line $-\sigma_0$ is to be prevented. Zeroing of OF_{13} will increase the damping further.

OF_{14} = an arbitrarily chosen very high fixed value (say, 10^6), which will indicate some σ_i values $\geq$ 0.0. This means unstable oscillation occurs for the particular parameters of PSS. These particular PSS parameters will be rejected during the optimization technique.

So, first multi-objective optimization function is formulated as in (2).

$$OF_1(\ ) = 10 \times OF_{11} + 10 \times OF_{12} + 0.01 \times OF_{13} + OF_{14} \tag{2}$$

The weighting factors '10' and '0.01' in (2) are chosen to impart more weights to OF_{11}, OF_{12} and to reduce high value of OF_{13}, to make them mutually competitive during optimization. By optimizing $OF_1(\ )$, closed loop system poles are consistently pushed further left of $j\omega$ axis with simultaneous reduction in imaginary parts also, thus, enhancing the relative stability and increasing the damping ratio above ξ_0. Finally, all closed loop system poles should lie within a D-shaped sector (Fig. 8) in the negative half plane of jω axis for which $\sigma_i << -\sigma_0$, $\xi_i >> \xi_0$. Selection of such low negative value of σ is purposefully chosen. The purpose is to push the closed loop system poles as much left as possible from the $j\omega$ axis to enhance stability to a great extent.

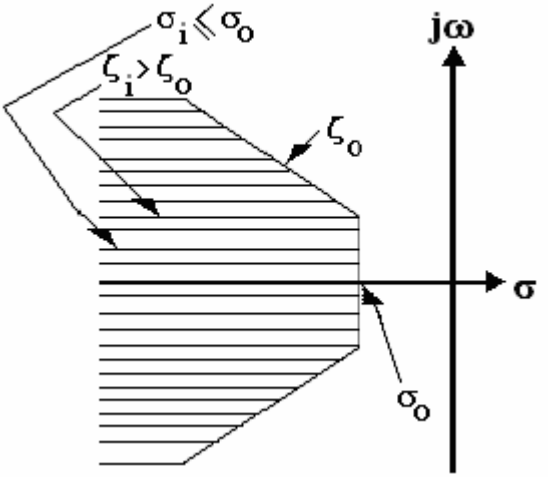


Fig. 8. D-shaped sector in the negative half of s plane

Thorough computation shows that optimization of $OF_1()$ is not sufficient for sharp tuning of PSS parameters. So, it is essential to design second multi-objective optimization function, $OF_2()$ for sharp tuning of PSS parameters. Thus, the second multi-objective optimization function is formulated as given in (3).

$$OF_2() = (o_{sh} \times 10^6)^2 + (u_{sh} \times 10^6)^2 + (t_{st})^2 + (\frac{d}{dt}(\Delta\omega_r) \times 10^6)^2 \quad (3)$$

In (3); o_{sh}, u_{sh}, t_{st}, $\frac{d}{dt}(\Delta\omega_r)$ are all referred to the transient response of $\Delta\omega_r$ determined by modal analysis subsequent to state perturbation $\Delta\delta = 5^\circ$ = 0.0857 rad. Results of subsequent optimizations are based on successive minimization of both $OF_1()$ and $OF_2()$, though most of the results show only $OF_1()$ values only, some show $OF_2()$ values only and some show both $OF_1()$ and $OF_2()$ values. The constrained optimization problem for the tuning of PSSs is, thus, formulated as follows:

Minimize $OF_1()$ and $OF_2()$ in succession with the help of any optimization technique to get optimal PSS parameters, subject to the limits [6, 8]:

a) For CPSS:

$$175.0 \le K_{pss} \le 230.0, \quad 0.001 \le T_{di} \le 1.0, \quad i = 1 \; to \; 6$$

b) For IEEE-PSS2B:

$$\begin{cases} 10 \le K_{s1} \le 30, \; 0.01 \le T_i \le 1.0, \quad i = 1 \; to \; 5, \\ T_{10} = 0.0001 \end{cases}$$

Other parameters are fixed.

c) For IEEE-PSS3B:

$$\begin{cases} -100.0 \le K_{s1} \le -10.0, 10.0 \le K_{s2} \le 100.0, \\ 0.005 \le T_{di} \le 2.0, \quad i = 1 \; to \; 4 \end{cases}$$

No fixed parameters are required.

d) For IEEE-PSS4B:

$$\begin{cases} -100.0 \le K_{s1} \le -10.0, 10.0 \le K_{s2} \le 100.0, \\ 0.005 \le T_i \le 2.0, \quad i = 1 \; to \; 4, \; T_0 = 0.2 \end{cases}$$

4 Tuning of PSS Parameters by Evolutionary Optimization Techniques

To improve the stabilizer's optimal performance and to achieve adaptive real time tuning and robustness under all sorts of system operating conditions and configurations, a lot of extensive research works have been reported in the literature. To recall a few among the other numerous works reported in the literature, PID PSS [10], pole-placement/shifting based PSS [11-12], artificial neural network PSS [13], H_2 optimal adaptive PSS [14], adaptive PSS [15], genetic local search [16], simulated annealing based PSS [17], design of sliding mode PSS via GA [18] may be noted.

4.1 GA for PSS Parameters Tuning

GA [19-21] is search procedures based on the mechanics of natural selection of genetics. Steps of GA, as implemented for optimization of PSS parameters, are presented in Fig. 9.

4.2 Bacteria Foraging Optimization

Natural selection favors propagation of genes of those animals that have efficient *foraging strategies* (method of finding, handling and taking in food) and eliminate those animals that have weak foraging strategies. As the efficient foraging strategy allows the animals to ingest better and quality food, only animals having better food searching strategy are allowed to enjoy reproductive cycle, in turn, producing better species. Poor foraging strategies are either shaped into good ones or eliminated after many generations. With their own physiological (e.g. cognitive and sensing capabilities) and environmental (e.g. physical characteristics of the search space, density of prey, risk and hazards from predators) constraints, animals try to maximize the consumption of energy per unit time interval. Such evolutionary idea has bred the concept of BFO [22] as an optimization algorithm. It is, gradually, being utilized by the interested research groups as an optimization algorithm to solve a range of non-linear optimization problems.

Four processes can explain the foraging strategy of *Escherichia* bacteria present in human intestine. These are *chemotaxis*, *swarming*, *reproduction* and *elimination-dispersal*.

A set of relatively rigid flagella helps the bacteria in locomotion. Its characteristics of movement for searching of food can be in two different ways, i.e. swimming and tumbling together known as *chemotaxis*. Its movement in a predefined direction is termed as *swimming* (*running*), whereas *tumbling* is the movement in

Step 1 Initialization:

a) Input operating values of P, Q, X_e, and E_t. Input fixed SMIB parameters,

b) Setting of limits of variable PSS parameters,

c) Setting of limits of GA parameters like mutation probability, crossover ratio,

d) Maximum population number of PSS parameter strings,

e) Maximum iteration cycles,

f) Binary value initialization of all the PSS parameter strings of the population,

g) Decoding of the binary strings within parameters' limits .

Step 2 Determination of SMIB parameters like $K_1 - K_6$

Step 3 Computation of objective function of each string of the total population

Step 4 Arraigning the values in increasing order from minimum and selection of top 50% better strings.

Step 5 Copying of 50% selected strings over the rest 50% inferior strings to form the total population

Step 6 Crossover

Step 7 Mutation

Step 8 Repeat from **Step 3** till the end of the maximum iteration cycle

Step 9 Determine the optimal PSS parameter string corresponding to the grand minimum misfitness.

Fig. 9. Steps of implementation of GA algorithm for tuning of PSS parameters

altogether different direction. During its entire lifetime, it alternates between these two modes of operation. Clockwise rotation of its flagella results in *tumble*, where as, anticlockwise rotation yields *swim*.

Likelihood of increased search for nutrient concentration, enhanced capability to gang up on a large prey to kill and digest it, group protection of the individual from predators are the highlighting objectives and motivations of social and intelligent foraging strategy. Successful foraging for food of each and individual of the group result from grouping, communication mechanism and collective intelligence. To attract the other bacteria towards the optimal convergence direction, it is necessary to pass on the information about the nutrient concentration (optimal point) to other bacteria. This is called *swarming*. To achieve this, a penalty function based upon the relative distances of each bacterium from the fittest bacterium

till that search duration, is added to the original optimization function. This penalty function becomes zero when all the bacteria have merged into the desired solution point.

After getting evolved through several chemotactic stages, the original set of bacteria is allowed for *reproduction*. Biological aspect of their conjugation process is splitting of one into two identical bacteria. It is mimicked in the optimization process by replacing the poorer half (having higher objective function value for minimization problem and vice versa) with weaker foraging strategy by the healthier half, which is eliminated owing to their poorer forging strategy, maintaining the total number of population bacteria constant in the process of evolution.

Elimination and/or *dispersal* of a set of bacteria to a new position result(s) in drastic alteration of smooth biological process of evolution. The underlying concept behind this step is to place a newer set of bacteria nearer to the food location to avoid *stagnation* (to avoid premature trapping into local optima). Pseudo code of BFO, as utilized for PSS parameter tuning, is depicted in Fig. 10.

Step 1 *Initialization:*

a) Maximum reproduction cycle, $\max_{reprod} = 10$,

b) Maximum chemotactic cycle, $\max_{chemo} = 20$,

c) Maximum dispersal cycle, $\max_{dispersal} = 1$,

d) Total number of bacteria, $numBact = 500$,

e) Maximum number of iteration cycle, (k), $\max_{cycles} = 200$

f) Some positive constants $d_{attract} = 2.0$; $w_{attract} = 0.2$; $d_{repelent} = 2.0$; $w_{repelent} = 0.1$; selection ratio, $s_r = 0.5$; probability of elimination, $P_{ed} = 0.3$; $c_{\max} = 0.1$; $c_{\min} = 0.0001$; $d_1 = 0.00001$; $d_2 = 0.00001$,

g) Maximum swim length (integer value) $\max_{swim} = 4$,

h) Number of problem variables, var (depends on specific application),

i) Variables' upper and lower limits $\text{var}_{\max}$ and $\text{var}_{\min}$, respectively, as given for individual test objective function.

Step 2 *Compute fitness functions of the bacteria,* J

Step 3 *Find pseudo global optimum fitness*

Step 4 *Find global optimum bacteria and its variables*

Step 5 *Swarming*

Fig. 10. Pseudo code of bacteria foraging optimization

$$J_{bact}^{(k)} = J_{bact}^{(k)} - \sum_{numBact} d_{attract} \times (e^{-w_{attarct} \times \sigma_{bact}^{(k)}}) + \sum_{numBact} h_{repellent} \times (e^{-w_{repellent} \times \sigma_{bact}^{(k)}}) \quad (4)$$

where,

$$\sigma_{bact}^{(k)} = \sigma_{bact}^{(k)} + \sum_{numBact} (\mathrm{var}_{global}^{(k)} - \mathrm{var}_{bact}^{(k)})^2 \quad (5)$$

Step 6 *Tumbling*

$$\Delta^{(\mathrm{var})} = (2 \times rand(\ rand()\ ) - 1) \times rand() \quad (6)$$

$$\Delta\Delta = \sqrt{\Delta(\mathrm{var}) \times \Delta^{'}(\mathrm{var})} \quad (7)$$

Compute:

$$\mathrm{var}^{(k+1)} = \mathrm{var}^{(k)} + (c_{\max} - d_2) \times \frac{\Delta^{(\mathrm{var})}}{\Delta\Delta} \quad (8)$$

Step 7 *Impose limits of variable*s

Step 8 *To start with BFO, the following pseudo codes are to be followed* :

(A) for disperse := 1 to $\max_{disperse}$ do

(B) for reprod := 1 to $\max_{reprod}$ do

(C) for chemo := 1 to $\max_{chemo}$ do

$$c^{(\mathrm{var})} = c_{\max}^{(\mathrm{var})} - (c_{\max}^{(\mathrm{var})} - c_{\min}^{(\mathrm{var})}) \times \frac{cycle}{\max cycles} \quad (9)$$

Step 9 *Compute fitness functions of the bacteria,* J

Step 10 *Find pseudo global optimum fitness,* J_{global}

Step 11 *Swarming: As shown in* ***Step 5***

Step 12 *Swarming of each bacteria*

(D) for bact := 1 to $numBact$ do

$m\_swim = 0$

Fig. 10. (*Continued*)

(E) while $m\_swim < swimlength$

$$m\_swim = m\_swim + 1$$

(F) if $J_{bact}^{(k)} < J_{bact}^{(k-1)}$

$xx = 'swim'$

Compute $\Delta^{(\mathrm{var})}$, $\Delta\Delta$ using (6) and (7)

Compute

$$\mathrm{var}^{(k)} = \mathrm{var}^{(k)} + (c^{(\mathrm{var})} + d_1^{(\mathrm{var})}{}_1) \times \frac{\Delta^{(\mathrm{var})}}{\Delta\Delta} \tag{10}$$

Impose restrictions on variables' values

Compute J_{bact} of each bacteria

Swarming: As shown in **Step 5**

else

$xx = 'Tumble'$

$J_{bact}^{(k)} < J_{bact}^{(k-1)}$

$m\_swim = swimlength$

Compute $\Delta^{(\mathrm{var})}$, $\Delta\Delta$ using (6) and (7)

Compute

$$\mathrm{var}^{(k)} = \mathrm{var}^{(k)} + (c^{(\mathrm{var})} - d_2^{(\mathrm{var})}) \times \frac{\Delta^{(\mathrm{var})}}{\Delta\Delta} \tag{11}$$

Impose restrictions on variables' values

Compute J_{bact} of each bacteria

Swarming: As shown in **Step 5**

End of if loop (F)

End of while loop (E)

End of for loop (D)

Fig. 10. (*Continued*)

Step 13 *Swarming: As shown in* ***Step 5***

Impose restrictions on variables' values

End of *chemo* loop (C)

Step 14 *Selection and Copying of elite bacteria*

Select elite bacteria by s_r ratio and copy the elite set of bacteria over the non-elite bacteria

Impose restrictions on variables' values

End of reprod loop (B)

End of disperse loop (A)

Step 15 Compute global optimum fitness among all previous pseudo global optimum fitness and plot.

Determine corresponding optimum problem variables.

Fig. 10. (*Continued*)

4.3 Chaotic Ant Swarm Optimization (CASO)

CASO [23] combines the chaotic and self-organization behavior of ants in the foraging process. It includes both effects of chaotic dynamics and swarm-based search. This algorithm is employed to tune the PSS (both single input and dual input) parameters with some modifications to suit the present application.

CASO is based on the chaotic behavior of individual ant and the intelligent organization actions of ant colony. Here, the search behavior of the single ant is "*chaotic*" at first and the organization variable, r_i is introduced to achieve self-organization process of the ant colony. Initially, the influence of the organization variable on the behavior of individual ant is sufficiently small. With the continual change of organization variable evolving in time and space, the chaotic behavior of the individual decreases gradually. Via the influence of the organization variable and the communication of previously best positions with neighbors, the individual ant alters his position and moves to the best one it can find in the search space.

The searching area of ants corresponds to the problem search space. In the search space R^l, which is the *l*-dimensional continuous space of real numbers, the algorithm searches for optima. A population of *K* ants is considered. These ants are located in a search space *S* and they try to minimize a function *f*: $S \rightarrow R$. Each point *s* in *S* is a valid solution to the considered problem. The position of an ant *i*

is assigned the algebraic variable symbol $S_i = (z_{i1},.....,z_{il})$, where i =1, 2, . . ., *K*. Naturally, each variable can be of any finite dimension. During its motion, each individual ant is influenced by the organization processes of the swarm. In mathematical terms, the strategy of movement of a single ant is assumed to be a function of the current position, the best position found by itself. Any member of its neighbors and the organization variable are given by (12).

$$\left.\begin{aligned} y_i(n) &= y_i(n-1)^{1+r_i} \\ z_{id}(n) &= \left(z_{id}(n-1) + \frac{c}{\psi_d} V_i \right) \times \\ &\exp\left((1-\exp(-a y_i(n))) \left(3 - \psi_d \left(z_{id}(n-1) + \frac{c}{\psi_d} V_i \right) \right) \right) \\ &- \frac{7.5}{\psi_d} V_i + \exp(-2a y_i(n) + b)(p_{id}(n-1) - z_{id}(n-1)) \end{aligned}\right\} \quad (12)$$

where

n	current iteration cycle,
$n-1$	previous iteration cycle,
$y_i(n)$	current state of the organization variable ($y_i(0) = 0.999$),
a , b , c	positive constants,
$r_i \in [0,0.1]$	a positive constant and is termed as the organization factor of ant *i*,
$z_{id}(n)$	current state of the d^{th} dimension of the individual ant *i*, $d = 1,2.....l$,
ψ_d	determines the selection of the search range of d^{th} element of variable in search space and
V_i	determines the search region of i^{th} ant and offers the advantage that ants could search diverse regions of the problem space ($V_i = rand()$).

The neighbor selection can be defined in the following two ways. The first one is the nearest fixed number of neighbors. The nearest *m* ants are defined as the neighbors of single i^{th} ant. The second way of the number of neighbor selection is to consider the situation in which the number of neighbors increases with iteration cycles. This is due to the influence of self-organization behaviors of ants. The impact of organization will become stronger than before and the neighbors of the ant

will increase. That is to say, the number of nearest neighbors is dynamically changed as iteration progresses.

The general CASO is a self-organizing system. When every individual trajectory is adjusted towards the successes of neighbors, the swarm converges or clusters in optimal regions of the search space. The search of some ants will fail if the individual cannot obtain information about the best food source from their neighbors.

For the work of this chapter, the algorithm's parameters, r_i, ψ_d, V_i, a, b, c are different from those as in [23]. These are respectively, $1+r_i$ is replaced by $(1.02+0.04\times rand())$, $\psi_d = 2.0$, $V_i = rand()$, $a=1$, $b=0.1$, $c=3$. These values are pre-set after a lot of experiments to get the best convergence to optimal solution. In our study, finally craziness is introduced as given in (13), which is not present in [23] but taken from [24, 25], introduced by Mukherjee and Ghoshal in the literature.

$$z_{id}(n) = z_{id}(n) + sign(r4)\times v_i^{craziness} \tag{13}$$

The values of $sign(r4)$ and $v_i^{craziness}$ are determined by (14) and (15), respectively.

$$sign(r4) = \begin{cases} 1, & (r4 \geq 0.5) \\ -1, & (r4 < 0.5) \end{cases} \tag{14}$$

$$v_i^{crazziness} = v_{\min}^{crazziness} + (v_{\max}^{crazziness} - v_{\min}^{crazziness})\times rand() \tag{15}$$

Introduction of craziness enhances CASO's ability of searching and convergence to a global optimal solution. Variables' upper and lower bound restrictions are always present. Ultimately, after maximum iteration cycles the optimal solution of z_{id} corresponds to the global optimal value of fitness function under consideration. Steps of implementation of CASO algorithm for tuning of PSS parameters are presented in Fig. 11.

5 Sugeno Fuzzy Logic as Applied for On-Line Tuning of PSS Parameters

The behavior of a synchronous generator connected to a network depends on, among other things, its position in this network, the operating conditions (in particular, the reactive power flow and the voltage map), the network topology and the generation schedule. Usage of the desired optimization technique yields a distinct set of controller parameters for different operating conditions. Under drastic change in operating conditions (e.g. different circuit topologies) the nominal controller is not necessarily going to be tuned enough to yield satisfactory performance. For on-line tuning of the parameters of PSS, very fast acting **Sugeno Fuzzy Logic** (SFL) is adopted.

Step 1 Initialization:

- Input operating values of P, Q, X_e, and E_t. Input fixed SMIB parameters,
- Setting of limits of variable PSS parameters,
- Setting of limits of CASO parameters,
- Maximum population number of PSS parameter strings, maximum iteration cycles,
- Real value initialization of all the PSS parameter strings of the population within limits.

Step 2 Determination of SMIB parameters like $K_1 - K_6$.

Step 3 Computation of misfitness function/objective function of each string of the total population.

Step 4 Determination of the best string corresponding to minimum misfitness/objective function value.

Step 5 Updating the strings of the population using CASO algorithm.

Step 6 Repeat from **Step 3** till the end of the maximum iteration cycle.

Step 7 Determine the optimal PSS parameter string corresponding to the grand minimum misfitness.

Fig. 11. Steps of implementation of CASO algorithm for tuning of PSS parameters

The whole process of SFL [20, 21] involves three steps as:

(a) Fuzzification: The first step is the fuzzification of input operating conditions as active power (P), reactive power (Q), equivalent transmission line reactance (X_e) and Generator's LT side bus voltage (E_t) in terms of fuzzy subsets (Low, Medium, High). These are associated with overlapping triangular membership functions. SFL rule base table is formed, each composed of four nominal inputs and corresponding nominal optimal PSS parameters as output determined by any of the optimizing techniques dealt with. The respective nominal central values of the input subsets of P are (0.2, 0.7, 1.2), those of Q are (-0.2, 0.6, 1.0), those of X_e are (0.4752, 0.77, 1.08) and those of E_t are (0.5, 0.8, 1.1), respectively, at which membership values are unity (Fig. 12). These are nominal input conditions also. Sugeno fuzzy rule base table consists of $3^4 (= 81)$ logical input conditions or sets (SFL tables calculated for different PSS structures investigated), each composed of four nominal inputs. Each logical input set corresponds to nominal optimal PSS parameters as output.

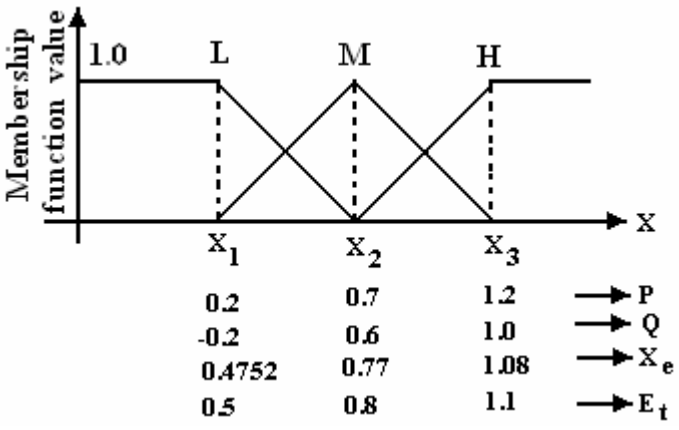


Fig. 12. Fuzzification of input operating conditions (P , Q , X_e and E_t)

(b) Sugeno Fuzzy Inference: For on-line imprecise values of input parameters, firstly their subsets in which the values lie are determined with the help of "IF", "THEN" logic and corresponding membership values are determined from the membership functions of the subsets. From Sugeno fuzzy rule base table, corresponding input sets and nominal PSS parameters are determined. Now, for each input set being satisfied, four membership values like μ_P, μ_Q, μ_{X_e} and μ_{E_t} and their minimum $\mu_{\min}$ are computed. For the input logical sets, which are not satisfied because parameters do not lie in the corresponding fuzzy subsets, $\mu_{\min}$ will be zero. For the non-zero $\mu_{\min}$ values only, nominal PSS parameters corresponding to fuzzy sets being satisfied are taken from the Sugeno fuzzy rule base table.

(c) Sugeno Defuzzification: It yields the defuzzified, crisp output for each parameter of PSS. Final crisp PSS parameter output is given by (16).

$$K_{crisp} = \frac{\sum_i \mu_{\min}^{(i)} . Ki}{\sum_i \mu_{\min}^{(i)}} \tag{16}$$

where i corresponds to input logical sets being satisfied among 81 input logical sets, K_i is the corresponding nominal PSS parameter. K_{crisp} is the crisp PSS parameter. $\mu_{\min}^{(i)}$ is the minimum membership value corresponding to the i^{th} input logical set being satisfied.

6 Input Parameters

(a) ***For SMIB System*:** Inertia constant, $H = 5$, $M = 2H$, nominal frequency, $f_0 = 50$ Hz, $0.995 \leq |\overline{E_b}| \leq 1.0$, the angle of $\overline{E_b} = 0^\circ$, $0.2 \leq P \leq 1.2$, $-0.2 \leq Q \leq 1.0$, $0.4752 \leq X_e \leq 1.08$, $0.5 \leq E_t \leq 1.1$. At t=0, it is assumed that $\Delta\delta = 5^\circ$ = 0.0857 rad is the state perturbation, $\Delta\omega_r = 0.0$ and other state deviations are also zero [6]. In the block diagram representation of generator with exciter and AVR, terminal voltage transducer time constant, $T_{rr} = 0.02\ s$, thyristor exciter gain, $K_a = 200.0$.

(b) ***For GA*:** Number of parameters depends on problem variables (PSS configuration), number of bits = (number of parameters)*8, population size = 50, maximum number of iteration cycles = 200, mutation probability = 0.001, crossover rate = 80%.

(c) ***For BFO*:** Number of problem variables depends upon the PSS structure under investigation. All the input parameters of the algorithm are given in Fig. 10.

(d) ***For CASO*:** Number of problem variables depends upon the PSS structure under investigation. All the input parameters of the algorithm are given in Section 4.3.

7 Performance Study of PSSs

Optimized PSS parameters determined by any of the optimization technique are substituted in MATLAB-SIMULINK model of the system to obtain the transient response profiles. Final values of $OF_1()$ and $OF_2()$, final eigenvalue, final undamped and damped frequencies and final damping ratio are all determined by the optimization technique at the end of optimization. Sugeno fuzzy rule base tables (not shown) are obtained by applying each optimization technique for distinct 81 number nominal input operating conditions. The outputs are 81 distinct nominal optimal PSS parameters sets. For the optimization, GA/BFO/CASO technique is adopted. GA is utilized for the sake of comparison.

7.1 Discussions on BFO-Based Results

a) **Analytical transient response characteristics:** Table 1 depicts the comparative GA and BFO based optimal transient response characteristics (in terms of $OF_2()$ value) of different PSS equipped system model. From this

table, it may be inferred that the transient stabilization performance of dual-input PSS equipped system model is better than single-input counterpart. Comparing dual-input PSSs, it is also observed that the transient stabilization performance of PSS3B equipped system model is superior to that of others. PSS3B equipped system model offers lesser values of o_{sh}, u_{sh}, t_{st}, $\frac{d}{dt}(\Delta\omega_r)$ and, thereby, lesser values of $OF_2(\,)$. It may also be observed that the BFO based optimization technique offer lesser value of $OF_2(\,)$ than GA based one. Thus, BFO based optimization technique offers better results than GA based one. Fig. 13 depicts that the comparative optimal transient performance of the different PSS equipped power system model corresponding to an operating condition of $P=1.2$, $Q=0.6$, $X_e=0.4752$, $E_t=0.5$ (all are in pu) for 0.01 pu simultaneous change in ΔT_m and ΔV_{ref}. From this figure, it is noticed that the transient stabilization performance of dual-input PSS is better than that of single input one. Among the dual-input PSSs, the performance of PSS3B is established to be the best one.

Table 1. GA and BFO based comparison of $OF_2(\,)$ values for CPSS, PSS2B, PSS3B and PSS4B based systems

Sl. No.	Per unit operating conditions (P, Q, X_e, E_t)	Algorithms	Value of $OF_2(\,)$ (X10^7)			
			CPSS	PSS2B	PSS3B	PSS4B
1	0.2, -0.2, 0.4752,1.1	GA-SFL	7.45	6.23	1.42	2.90
		BFO-SFL	4.91	4.02	**1.01**	1.62
2	0.5, 0.2, 0.4752,1.0	GA-SFL	7.85	7.33	2.47	3.17
		BFO-SFL	4.87	4.14	**1.21**	1.52
3	0.75, 0.50, 0.4752, 0.5	GA-SFL	7.16	5.18	2.71	3.88
		BFO-SFL	3.97	3.17	**1.01**	1.45
4	0.95, 0.30, 0.4752, 0.5	GA-SFL	7.29	5.85	2.02	4.58
		BFO-SFL	3.15	2.17	**1.42**	1.96
5	1.2, 0.6, 1.08, 0.5	GA-SFL	8.96	8.72	3.74	4.39
		BFO-SFL	5.16	4.83	**2.57**	2.43

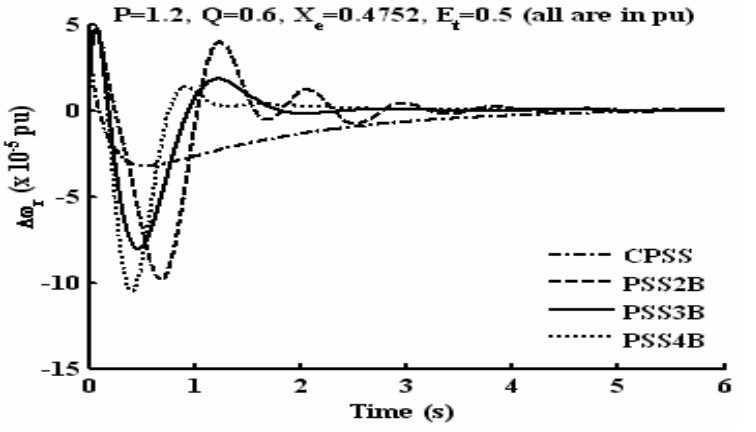


Fig. 13. Comparative BFO based transient response profiles of CPSS, PSS2B, PSS3B and PSS4B for 0.01 pu simultaneous change in ΔT_m and ΔV_{ref}

b) **Analytical eigenvalue based system performance analysis:** BFO based comparison of $OF_1()$ values of CPSS, PSS2B, PSS3B and PSS4B are shown in Table 2 for different system operating conditions. From Table 2, it is observed that the value of $OF_1()$ is the least for PSS3B, establishing the performance of PSS3B to be the best one. For PSS3B equipped system model, majority of the eigenvalues are within D-shaped sector (Fig. 8) which yield lesser values of $OF_{11}()$, $OF_{12}()$ and $OF_{13}()$. Hence, the value of $OF_1()$ is very less for PSS3B equipped system model. On the other hand, majority of the eigenvalues for PSS2B based system are outside the D-shaped sector but very close to and at the right side of $(-\sigma_0, j0)$ point. This yields higher values of $OF_{11}()$, $OF_{12}()$ and $OF_{13}()$. The value of $OF_1()$ is more for this system. Thus, from the eigenvalue analysis it may be concluded that a considerable improvement has occurred in the transient performance for the PSS3B based system.

c) **Convergence profiles:** The comparative BFO based convergence profiles of $OF_1()$ for CPSS, PSS2B, PSS3B and PSS4B are depicted in Fig. 14 corresponding to an operating condition of $P = 0.5$, $Q = 0.2$, $X_e = 0.4752$, $E_t = 1.0$ (all are in pu). From these figures, the objective function value, $OF_1()$ corresponding to PSS3B is found to converge faster than the others.

Table 2. BFO based comparison of $OF_1()$ values under different operating conditions for CPSS, PSS2B, PSS3B and PSS4B

Operating conditions	Type of PSS	PSS parameters	$OF_1()$	t_{ex} (s)
1.2, 0.6, 0.4752, 0.5	CPSS	175.00, 0.005, 0.005, 0.001, 0.001, 0.352, 0.001	1455.62	330.65
	PSS2B	10.00, 0.097, 0.010, 0.113, 0.01, 0.25	1251.60	639.21
	PSS3B	**-10.00, 10.00, 0.271, 0.005, 0.005, 0.334**	**291.30**	344.78
	PSS4B	-10.00, 10.00, 0.273, 0.005, 0.138, 0.005	311.52	420.73
1.0, 0.2, 1.08, 0.5	CPSS	230.00, 0.005, 0.005, 0.029, 0.001, 0.001, 0.001	1491.69	345.16
	PSS2B	10.00, 0.01, 0.126, 0.01, 0.393, 0.204	1444.20	638.75
	PSS3B	**-10.00, 10.00, 0.171, 0.005, 0.139, 0.131**	**267.77**	335.89
	PSS4B	-10.00, 10.00, 0.005, 0.168, 0.237, 0.005	372.49	415.62

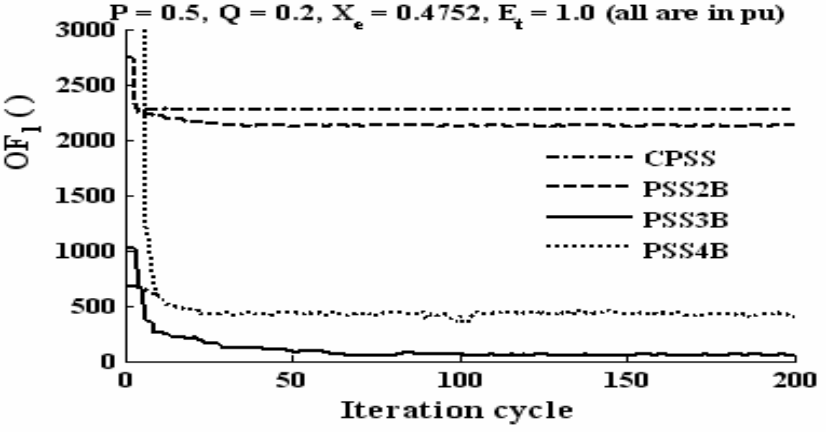


Fig. 14. Comparative BFO based convergence profiles of $OF_1()$ for CPSS, PSS2B, PSS3B and PSS4B ($P = 0.5$, $Q = 0.2$, $X_e = 0.4752$, $E_t = 1.0$)

d) Comparative optimization performance of the optimization techniques: With regard to optimization performances of the optimizing algorithms, as depicted in Table 1, it may be concluded that the BFO based approach offers the lower values of $OF_2()$ for the same input operating

conditions. Comparative transient performances of $\Delta\omega_r$ and convergence profiles of $OF_1()$ for BFO and GA based optimization for all the four PSS modules(CPSS, PSS2B, PSS3B, and PSS4B) are depicted in Fig. 15 and Fig. 16 respectively. These figures assist to conclude that the transient stabilization performance and convergence profile of BFO based optimization are better than those of GA based one. Though the execution time of BFO (Table 2) is more than GA (not shown in the present work), BFO offers much better optimal performance. Thus, BFO may be accepted as a better optimizing algorithm.

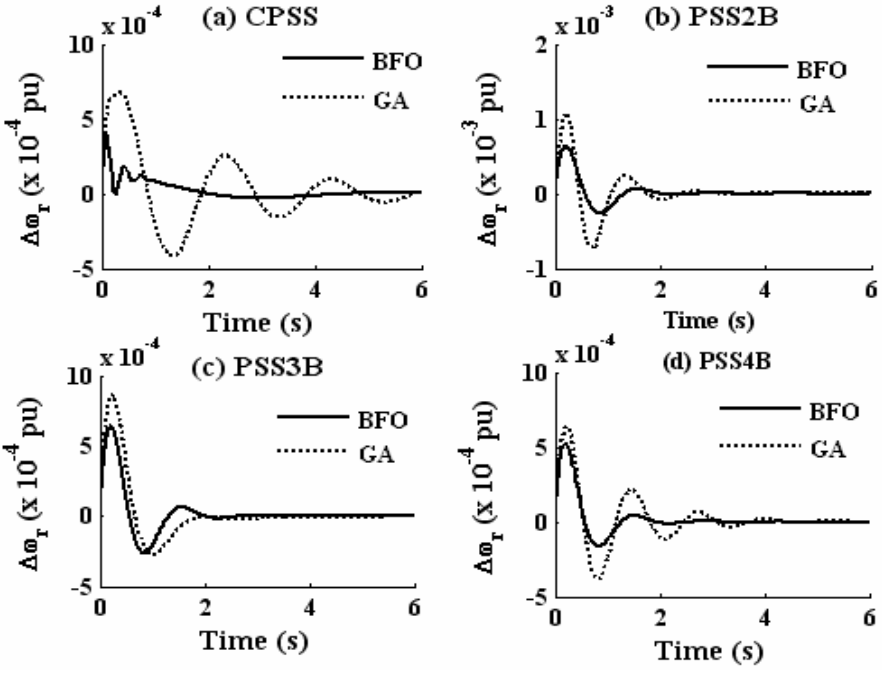


Fig. 15. Comparative BFO and GA based transient response profiles of CPSS, PSS2B, PSS3B and PSS4B for 0.01 pu simultaneous change in ΔT_m and ΔV_{ref} ; ($P = 0.2$, $Q = 0.2$, $X_e = 0.4752$, $E_t = 1.0$)

e) Transient performance study under different perturbations: Fig. 17 displays MATLAB-SIMULINK based transient response profiles of $\Delta\omega_r$ and Δv_1 for PSS3B equipped generator. This figure helps to conclude that PSS3B damps the oscillations of $\Delta\omega_r$ and Δv_1 very quickly under any form of system perturbations. Real parts of some eigenvalues for CPSS / PSS2B are always either equal to, or greater than σ_0 in the

negative half plane of $j\omega$ axis. A few eigenvalues are always outside D-shaped sector (Fig. 8) for any operating condition. So, objective function values ($OF_1()$) are always higher (Table 2). Much lower negative real parts of eigenvalues of PSS3B and PSS4B (not shown) cause higher relative stability than CPSS / PSS2B. Larger reductions of ω_n and ω_d for some electromehanical oscillations are due to higher damping ratios $\xi_i >> \xi_0$ for those particular modes, in case of PSS3B and PSS4B (Table 3).

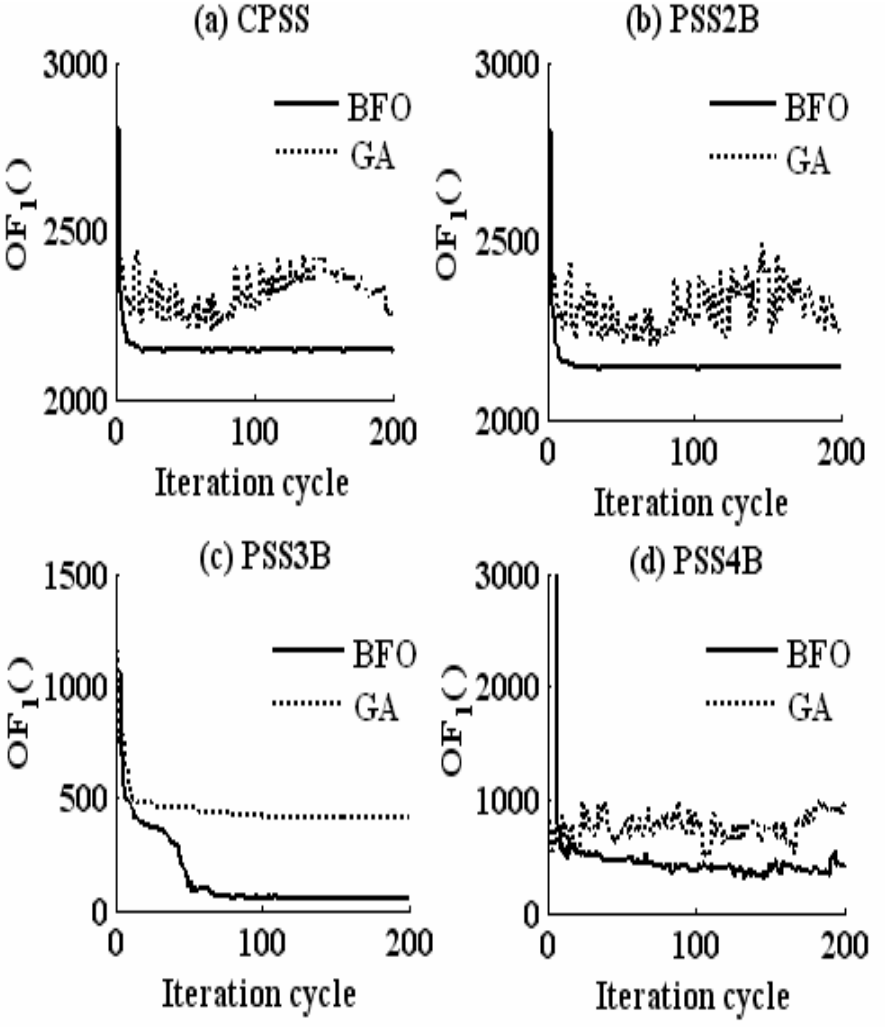


Fig. 16. Comparative BFO and GA based transient response profiles of CPSS, PSS2B, PSS3B and PSS4B for 0.01 pu simultaneous change in ΔT_m and ΔV_{ref}; ($P = 0.2$, $Q = 0.2$, $X_e = 0.4752$, $E_t = 1.0$)

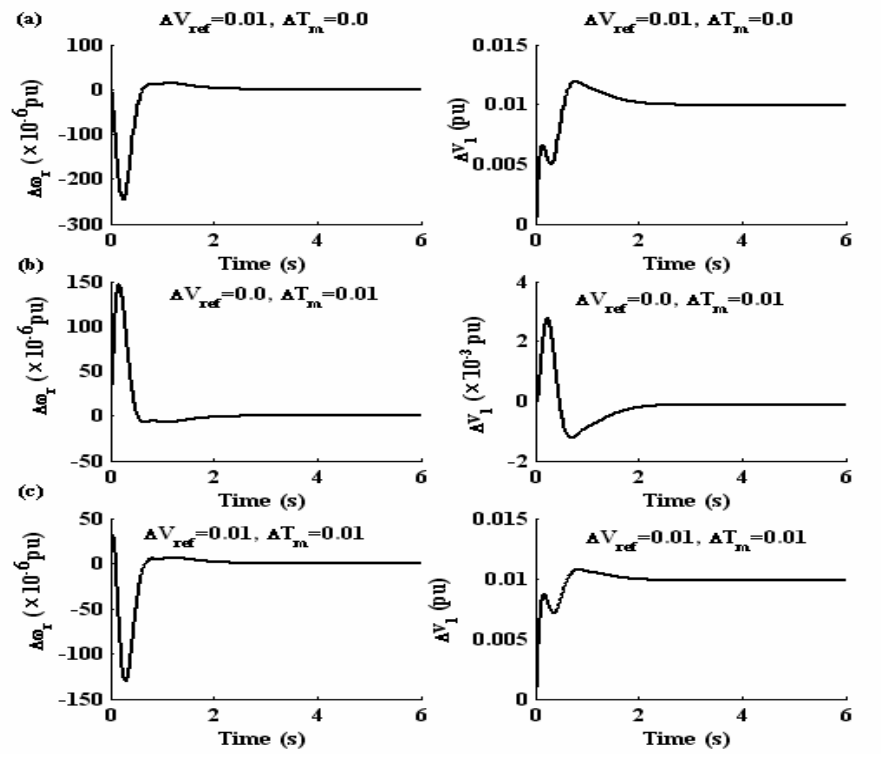


Fig. 17. MATLAB-SIMULINK based transient response profiles of $\Delta\omega_r$ and Δv_1 under different perturbation conditions for generator equipped with PSS3B

Table 3. GA, and BFO based results of eigenvalue analysis corresponding to operating conditions ($P = 0.95$, $Q = 0.30$, $X_e = 1.08$, $E_t = 0.5$; all are in pu)

Type of PSS	Algorithms-SFL	Damping ratio (ξ)		Undamped natural frequency (ω_n), rad/s		Corresponding damped frequency (ω_d), rad/s	
		L*	H*	L*	H*	L*	H*
CPSS	I#	0.16	0.57	0.38	2.45	0.42	0.59
	II#	0.38	0.65	0.59	1.45	0.45	4.31
PSS2B	I#	0.26	0.97	0.48	3.33	0.32	0.79
	II#	0.49	1.00	0.63	13.42	0.56	7.79
PSS3B	I#	0.36	0.98	0.17	1.93	0.15	1.8
	II#	**0.72**	**0.98**	**0.47**	**0.96**	**0.33**	**0.38**
PSS4B	I#	0.2	0.95	0.55	1.22	0.54	0.36
	II#	0.37	0.95	0.34	2.97	0.28	2.98

I# means GA-SFL, II# means BFO-SFL

L* means lowest, H* means highest.

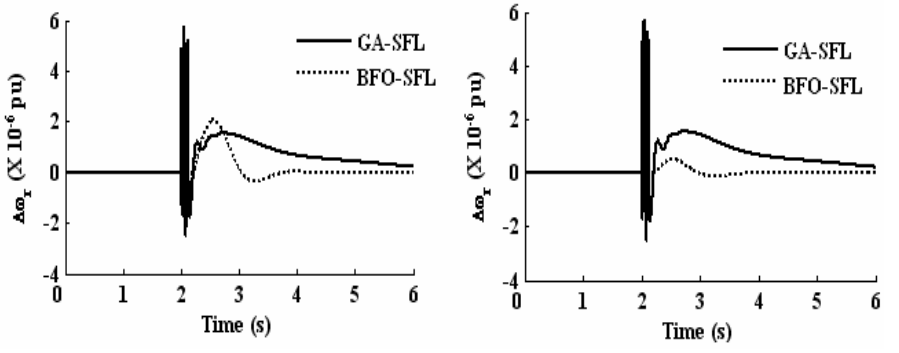


Fig. 18. Comparative GA-SFL and BFO-SFL based transient response profiles of $\Delta\omega_r$ for the generator equipped with PSS3B under change in operating conditions

Table 4. Off-nominal operating conditions, simulation of faults, algorithms-SFL, and optimal PSS3B parameters

Fault No.	Off-nominal operating conditions (P, Q, X_e, E_t ; all are in pu)	Algorithms-SFL	PSS3B parameters
Fault 1	0.95, 0.3, 0.4752, 1.0	BFO-SFL	-13.50, 83.26, 0.529, 0.051, 0.008, 1.566
		GA-SFL	-59.57, 59.22, 0.558, 0.013, 0.317, 0.169
	LT bus fault of duration 220 ms and subsequent clearing	BFO-SFL GA-SFL	No change in parameters
	0.95, 0.3, 1.08, 1.0	BFO-SFL	-12.34, 83.67, 1.608, 0.239, 0.326, 0.473
		GA-SFL	-14.22, 66.25, 0.208, 0.013, 0.161, 0.099
Fault 2	1.0, 0.6, 0.4752, 1.1	BFO-SFL	-10.88, 26.40, 1.279, 0.369, 0.070, 0.140
		GA-SFL	-10.00, 53.24, 1.618, 0.192, 0.410, 0.254
	LT bus fault of duration 220ms and subsequent clearing	BFO-SFL GA-SFL	No change in parameters
	0.2, 0.2, 0.4752, 1.1	BFO-SFL	-13.07, 39.00, 1.322, 1.816, 1.006, 0.088
		GA-SFL	-15.27, 12.11, 1.805, 0.893, 0.137, 0.005

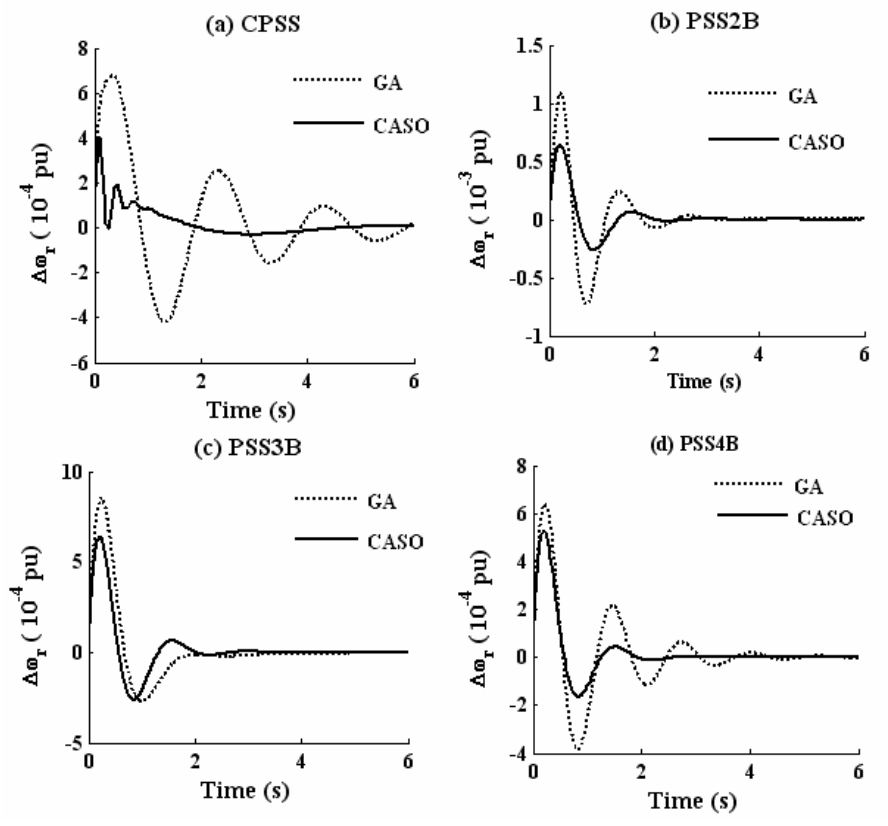


Fig. 19. Comparative CASO and GA-based transient response profiles of $\Delta\omega_r$ for CPSS, PSS2B, PSS3B and PSS4B for 0.01 pu simultaneous change in ΔT_m and ΔV_{ref} ($P = 1.0$, $Q = 0.2$, $X_e = 1.08$, $E_t = 0.9$, all are in pu)

f) Simulation of fault: It is revealed that PSS3B is offering the best transient performance in damping all electromechanical modes of generator's angular speed oscillations for all nominal and off-nominal system conditions, step changes of mechanical torque inputs (ΔT_m), reference voltage inputs (ΔV_{ref}). LT bus fault of duration 220 ms at the instant of 2.0 s is simulated for PSS3B equipped system model and the corresponding comparative transient response profiles of $\Delta\omega_r$ for both GA-SFL and BFO-SFL based responses are plotted in Fig. 18 (Fig. 18(a) corresponds to Fault 1 and Fig. 18(b) corresponds to Fault 2). A look into these figures show that after the creation of the fault the BFO-SFL based response recovers from this abnormal situation with much lesser fluctuation in angular speed as compared to that of GA-SFL based one.

Table 4 depicts the system model parameters as determined by SFL. Thus, BFO-SFL based model exhibits better response having lesser amplitude of angular speed deviation under fault and subsequent clearing condition yielding better dynamic robust transient performance than GA-SFL based one. Thus, PSS3B proves to be much less susceptible to faults because PSS3B settles all the state deviations to zero much faster than any other PSSs.

7.2 Discussions on CASO-Based Results

a) Comparative optimization performance of the optimization techniques: Comparative transient performances of $\Delta\omega_r$ and convergence profiles of $OF_1(\,)$ for GA and CASO-based optimization for all the four PSS modules (CPSS, PSS2B, PSS3B, and PSS4B) are depicted in Fig. 19 and Fig. 20,

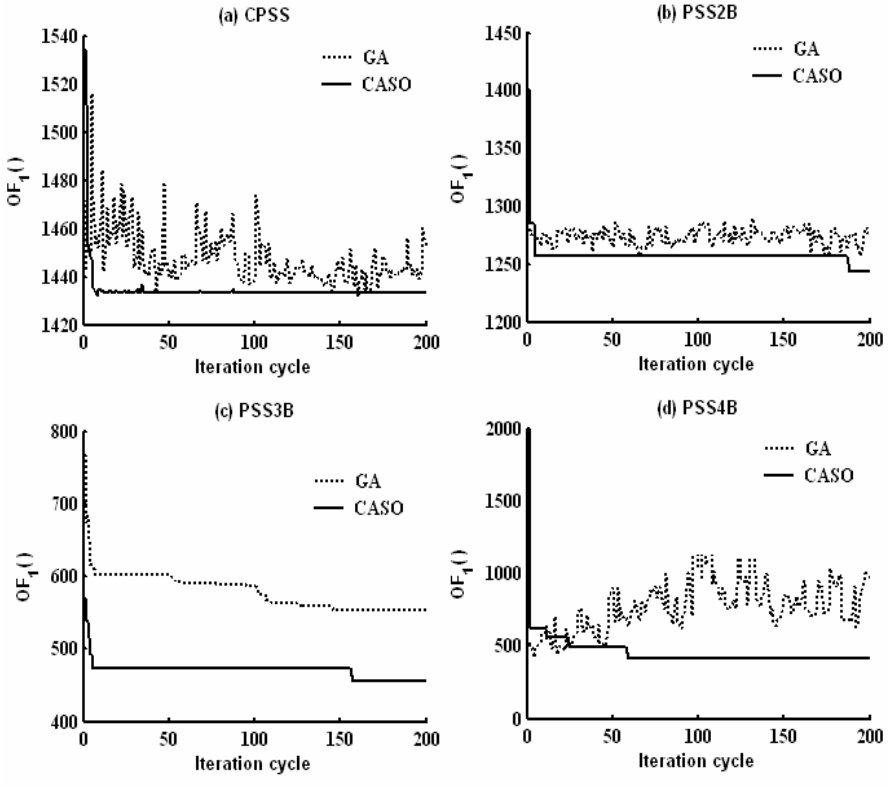


Fig. 20. Comparative CASO, and GA-based transient response profiles of $OF_1(\,)$ CPSS, PSS2B, PSS3B and PSS4B for 0.01 pu simultaneous change in ΔT_m and ΔV_{ref} ($P = 1.0$, $Q = 0.2$, $X_e = 1.08$, $E_t = 0.9$, all are in pu)

respectively. From Fig. 19, it may be concluded that the transient stabilization performance of CASO-based optimization is better than that of GA-based one. From Fig. 20, it is observed that CASO-based optimization technique offers lesser $OF_1(\,)$ value as compared to GA-based one, for whatever may be the PSS model under consideration. Thus, CASO offers much better optimal performance.

b) Fuzzy logic-based tuning of PSS parameters under changes in system operating conditions: LT bus fault of duration 220 ms at the instant of 2.0 s is simulated for different PSS equipped system model and the corresponding comparative transient response profiles of $\Delta\omega_r$ for both GA-SFL and CASO-SFL-based approaches are plotted in Fig. 21 for all the PSSs.

Fig. 21 reveals that CASO-SFL response exhibits better response than GA-SFL based one irrespective of the PSS model. A close look into this figure also reveals that after the creation of the fault, the CASO-SFL-based response for

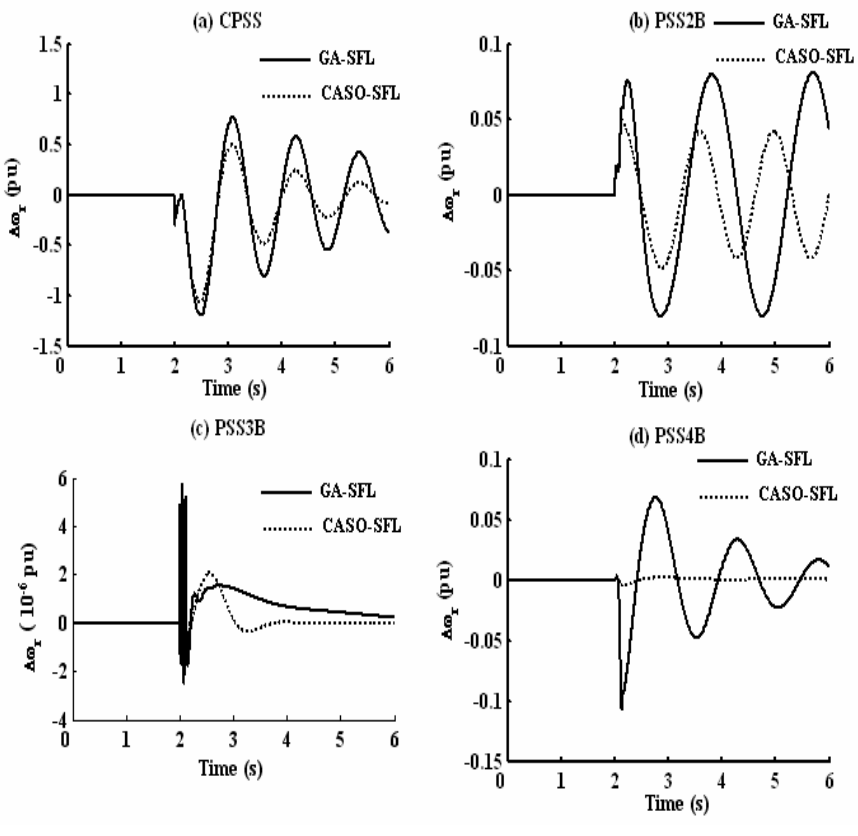


Fig. 21. Comparative GA-SFL, and CASO-SFL-based transient response profiles of $\Delta\omega_r$ for the generator equipped with PSS under change in operating conditions

Table 5. Off-nominal operating conditions, type of PSS, algorithms-SFL, and optimal PSS parameters

Off-nominal operating conditions	Type of PSS	Algorithms-SFL	PSS parameters
1.0, 0.5, 0.4752, 1.0 (Pre-fault)	CPSS	GA-SFL	179.68, 0.030, 0.032, 0.245, 0.005, 0.144, 0.401
		CASO-SFL	230.00, 0.177, 0.005, 0.001, 0.001, 0.001, 0.001
	PSS2B	GA-SFL	10.00, 0.143, 0.047, 0.089, 0.011, 0.138
		CASO-SFL	10.00, 0.098, 0.010, 0.010, 0.039, 0.022
	PSS3B	GA-SFL	-10.00, 83.47, 1.992, 0.995, 0.005, 0.036
		CASO-SFL	-10.00, 10.00, 2.00, 1.142, 0.005, 0.077
	PSS4B	GA-SFL	-27.23, 55.70, 0.440, 0.005, 0.044, 0.005
		CASO-SFL	-10.00, 10.00, 0.579, 0.005, 0.202, 0.005
LT bus fault of duration 220 ms and subsequent clearing	CPSS/ PSS2B/ PSS3B/ PSS4B	GA-SFL CASO-SFL	No change in parameters
0.2, -0.2, 1.08, 1.1 (Post-fault)	CPSS	GA-SFL	182.32, 0.109, 0.090, 0.071, 0.038, 0.088, 0.075
		CASO-SFL	175.00, 0.065, 0.005, 0.001, 0.001, 0.001, 0.001
	PSS2B	GA-SFL	10.39, 0.022, 0.150, 0.177, 0.179, 0.077
		CASO-SFL	10.00, 0.282, 0.010, 0.795, 0.010, 0.010
	PSS3B	GA-SFL	-18.09, 10.00, 1.525, 0.994, 0.005, 0.005
		CASO-SFL	-10.00, 10.00, 0.861, 2.00, 0.005, 0.005
	PSS4B	GA-SFL	-10.35, 69.06, 0.153, 0.005, 0.852, 0.005
		CASO-SFL	-10.00, 10.00, 0.139, 0.005, 0.076, 0.014

PSS3B equipped system model recovers from this abnormal situation with much lesser fluctuation in angular speed as compared to that of GA-SFL-based counter part.

Table 5 depicts the system model parameters as determined by SFL for all the PSS models. Thus, CASO-SFL-based model exhibits better response having lesser amplitude of angular speed deviation under fault and subsequent clearing condition yielding better dynamic transient performance than GA-SFL-based one for PSS3B equipped system model. Hence, PSS3B proves to be much less susceptible to faults because PSS3B settles all the state deviations to zero much faster than any other PSS.

c) Comparative performances of PSSs in terms of analytical transient response characteristics: From Table 6, it may also be inferred that the transient stabilization performance of dual-input PSS equipped system model is better than single-input counter part. Comparing dual-input PSSs, it is also observed that the transient stabilization performance of PSS3B equipped system model is superior to that of others. PSS3B equipped system model offers lesser values of o_{sh}, u_{sh}, t_{st}, $\frac{d}{dt}(\Delta\omega_r)$ and, thereby, lesser values of $OF_2(\,)$. Fig. 22 depicts that the comparative optimal transient performance of the different PSS equipped power system model corresponding to an operating condition of $P = 0.5$, $Q = 0.2$, $X_e = 0.93$, $E_t = 1.0$ (all are in pu) for 0.01 pu simultaneous change in ΔT_m and ΔV_{ref}. From this figure, it is noticed that the transient stabilization performance of dual-input PSS is better than that of single input one. Among the dual-input PSSs, the performance of PSS3B is established to be the best one for this specific application of SMIB system model under study. From Fig. 22, it is apparently observed that CPSS outperforms the other three PSSs in terms of magnitudes of u_{sh} and o_{sh}. But the duration of undershoot and t_{st} are very large for CPSS. The second objective function $OF_2(\,)$ is designed in such a fashion that it will take care of o_{sh}, u_{sh}, t_{st}, $\frac{d}{dt}(\Delta\omega_r)$. And as the value of t_{st} is very large for CPSS, it offers the highest value of $OF_2(\,)$ among all the PSSs taken in this study. Thus, CPSS exhibits poor performance for the specific application under study.

Table 6. GA, and CASO-based comparison of $OF_2(\,)$ values for CPSS, PSS2B, PSS3B and PSS4B-based systems

Sl. No.	Operating conditions	Algorithms	Value of $OF_2(\,)$ (X10^7)			
			CPSS	PSS2B	PSS3B	PSS4B
1	0.2, -0.2,	GA-SFL	7.45	6.23	1.42	2.90
	0.4752,1.1	CASO-SFL	**3.24**	**3.97**	**0.99**	**1.47**
2	0.5, 0.2,	GA-SFL	7.85	7.33	2.47	3.17
	0.4752, 1.0	CASO -SFL	**4.19**	**4.04**	**1.12**	**1.15**
3	0.75, 0.50,	GA-SFL	7.16	5.18	2.71	3.88
	0.4752, 0.50	CASO -SFL	**3.14**	**2.47**	**0.87**	**1.19**
4	0.95, 0.30,	GA-SFL	7.29	5.85	2.02	4.58
	0.4752, 0.5	CASO-SFL	**3.01**	**2.06**	**1.15**	**1.45**
5	1.2, 0.6,	GA-SFL	8.96	8.72	3.74	4.39
	1.08, 0.5	CASO-SFL	**4.19**	**3.89**	**2.01**	**2.01**

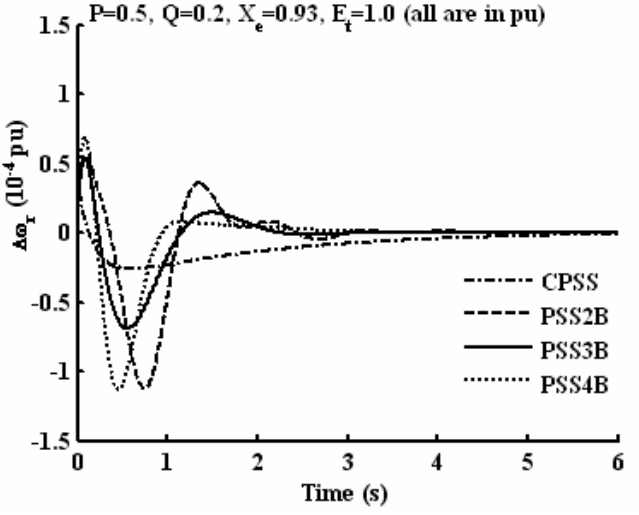


Fig. 22. Comparative CASO-based transient response profiles of $\Delta\omega_r$ for CPSS, PSS2B, PSS3B and PSS4B for 0.01 pu simultaneous change in ΔT_m and ΔV_{ref} ($P = 0.2$, $Q = 0.2$, $X_e = 0.93$, $E_t = 1.0$, all are in pu)

d) Comparative performances of PSSs based on convergence profiles: The comparative CASO-based convergence profiles of $OF_1(\,)$ for CPSS, PSS2B, PSS3B, and PSS4B are depicted in Fig. 23 corresponding to an operating condition of $P = 1.0$, $Q = 0.2$, $X_e = 0.4752$, $E_t = 1.0$ (all are in pu). From these figures, the objective function value, $OF_1(\,)$ corresponding to PSS3B is found to be the least one and, hence, PSS3B probes to be the best PSS among the PSSs considered for the specific application under study.

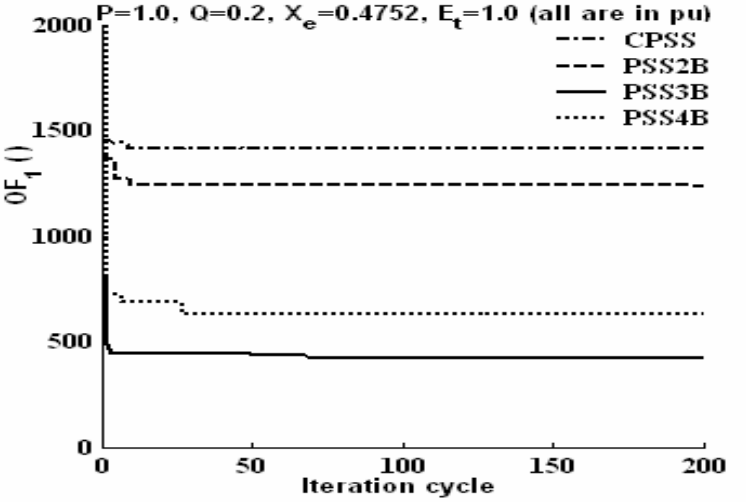


Fig. 23. Comparative CASO-based convergence profiles of $OF_1(\,)$ for CPSS, PSS2B, PSS3B and PSS4B ($P = 1.0$, $Q = 0.2$, $X_e = 0.4752$, $E_t = 1.0$, all are in pu)

e) Performance improvement of CASO with craziness: In this chapter, the concept of craziness (a novel concept proposed by Mukherjee and Ghoshal in the literature [24, 25]) is blended with original CASO algorithm with an attempt to have improved performance of CASO algorithm. Table 7 shows how improved performance in CASO is taking place with the inclusion of craziness concept.

Table 7. Comparison of $OF_1(\,)$ values based on CASO without craziness and CASO with craziness for all the PSSs, nominal input operating conditions being $P = 0.2$, $Q = 0.2$, $X_e = 1.08$, $E_t = 1.0$; all are in pu

Type of PSS	Type of Algorithm	PSS parameters	$OF_1(\,)$
CPSS	Algo 1*	175.00, 0.734, 1.00, 0.001, 0.001, 0.001, 0.001	1375.53
	Algo 2#	**175.00, 0.082, 0.005, 0.001, 0.001, 0.001, 0.001**	**1299.39**
PSS2B	Algo 1*	10.00, 0.375, 0.010, 1.00, 0.010, 0.010	1227.92
	Algo 2#	**10.00, 1.00, 0.011, 0.253, 0.010, 0.010**	**1200.61**
PSS3B	Algo 1*	-10.00, 10.00, 0.965, 1.510, 0.040, 0.005	449.23
	Algo 2#	**-10.00, 10.00, 0.934, 2.00, 0.012, 0.005**	**442.41**
PSS4B	Algo 1*	-10.00, 10.00, 0.229, 0.202, 0.191, 0.019	560.42
	Algo 2#	**-10.00, 10.00, 0.271, 0.103, 0.005, 0.017**	**518.35**

Algo 1* means CASO without craziness, Algo 2# CASO with craziness.

This table shows for a particular input operating condition, improved $OF_1()$ value is achieved for all the four PSS structures. The comparative convergence profiles of $OF_1()$ based on CASO without craziness and CASO with craziness for CPSS, PSS2B, PSS3B, and PSS4B are depicted in Fig. 24 corresponding to a nominal operating condition of $P = 0.2$,

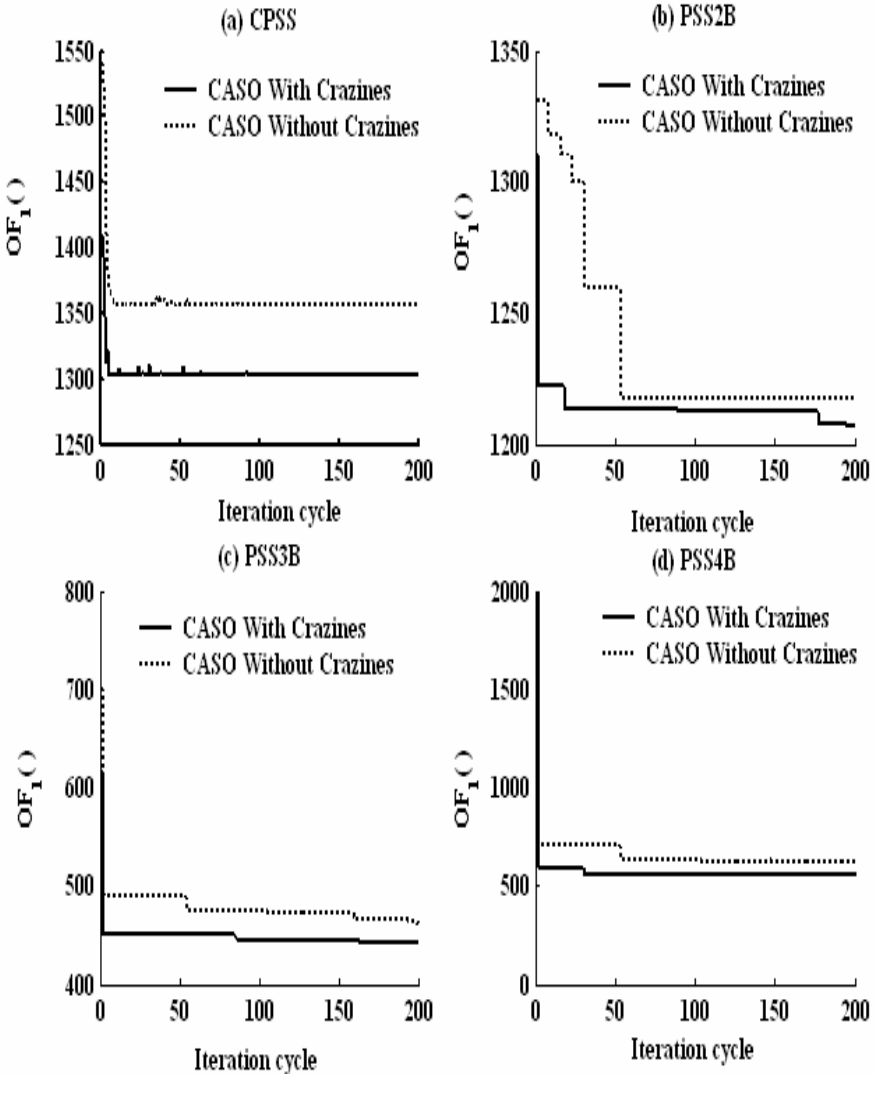


Fig. 24. Comparative CASO (with craziness and without craziness)-based convergence profiles of $OF_1()$ for CPSS, PSS2B, PSS3B and PSS4B ($P = 0.2$, $Q = 0.2$, $X_e = 1.08$, $E_t = 1.0$, all are in pu)

$Q = 0.2$, $X_e = 1.08$, $E_t = 1.0$ (all are in pu). From this figure, the objective function value, $OF_1()$ corresponding to CASO with craziness is found to converge to the lesser minimum value faster than the others. It is also noticed that the $OF_1()$ value of PSS3B corresponding to CASO with craziness is the lowest one.

f) **System performance for generator without any PSS:** Fig. 25 exhibits MATLAB-SIMULINK based plot of the transient performance of $\Delta\omega_r$ for simultaneous change in ΔT_m and ΔV_{ref} in case of generator without any PSS. The dynamic response of the system without any PSS is unstable even due to any small change in mechanical input torque or reference voltage. So, it justifies the application of PSS, in general, to enhance small-signal stability, especially, in case of synchronous generator equipped with very high gain static excitation systems.

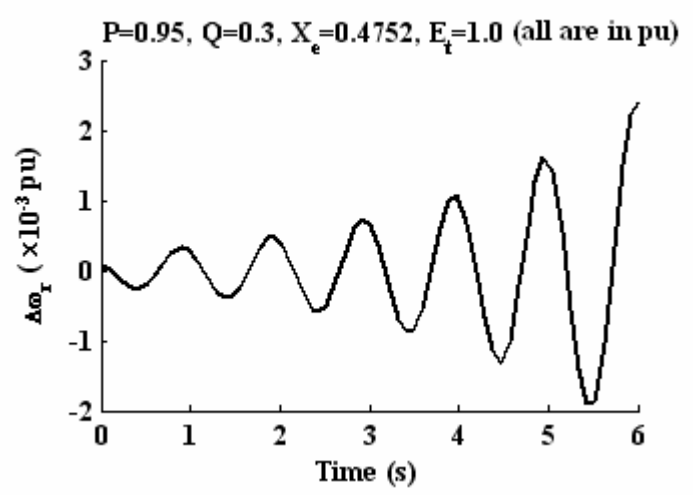


Fig. 25. MATLAB-SIMULINK based plot of $\Delta\omega_r$ for simultaneous change in ΔT_m and ΔV_{ref} in case of generator without PSS

8 Classical Phase Compensation Based PSS vis-à-vis IEEE Dual Input PSS

Besides comparison of several evolutionary techniques applied for optimal phase compensation with three types of IEEE dual input PSSs, classical phase compensation based conventional PSS (CPSS) is presented for the purpose of benchmarking in this chapter. The speed deviation response profile for the SMIB system, as considered in this chapter, with CPSS/PSS2B/PSS3B/PSS4B

(with very high gain exciter, $K_A = 200$) and without PSS (with low gain exciter, $K_A = 5$) for $P = 1.2$, $Q = 0.6$, $X_e = 0.93$, $E_t = 1.0$ (all are in pu) and 0.01 step perturbation in ΔT_m and no change in ΔV_{ref} is presented in Fig. 26. From this figure, it is inferred that the dual input PSS produces additional damping to settle the oscillations much quicker than CPSS. For the system without PSS, $K_A = 5$ is chosen instead of choosing $K_A = 200$ because high gain exciter without PSS causes instability in the system as illustrated in Fig. 25.

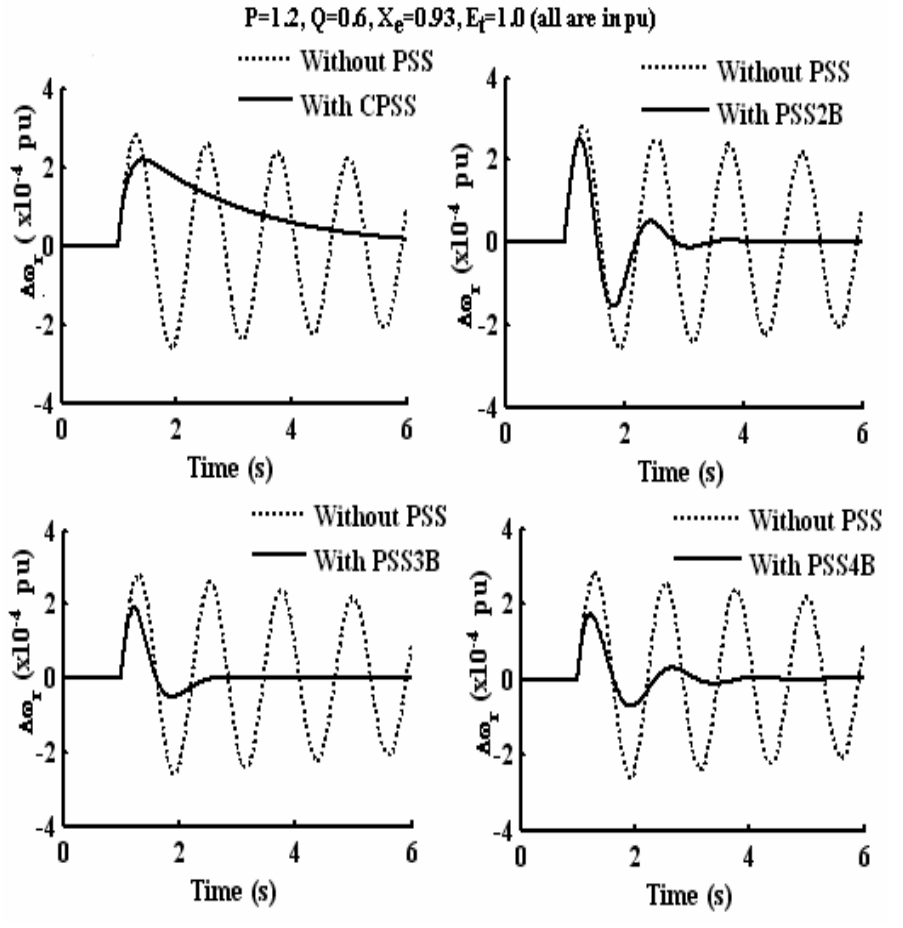


Fig. 26. Comparative damping performances of different PSS models for an SMIB model under 0.01 pu step change in ΔT_m and no change in ΔV_{ref} (with CASO based optimized PSS parameters)

The comparative degrees of phase compensation offered by different dual input IEEE PSSs are contrasted with that of conventional PSS and system without PSS in Fig. 27, for $P = 1.2$, $Q = 0.6$, $X_e = 0.93$, $E_t = 1.0$ (all are in pu) and 0.01 simultaneous step perturbation in ΔT_m and ΔV_{ref}. This figure reveals PSS3B offers maximum damping to settle down the frequency deviation. Thus the stabilizing action offered by PSS3B is the best one among the PSSs considered, at least for the SMIB system considered. Thus, a proper benchmarking against the conventional techniques – e.g. classical phase compensation based PSS is highlighted in the present chapter.

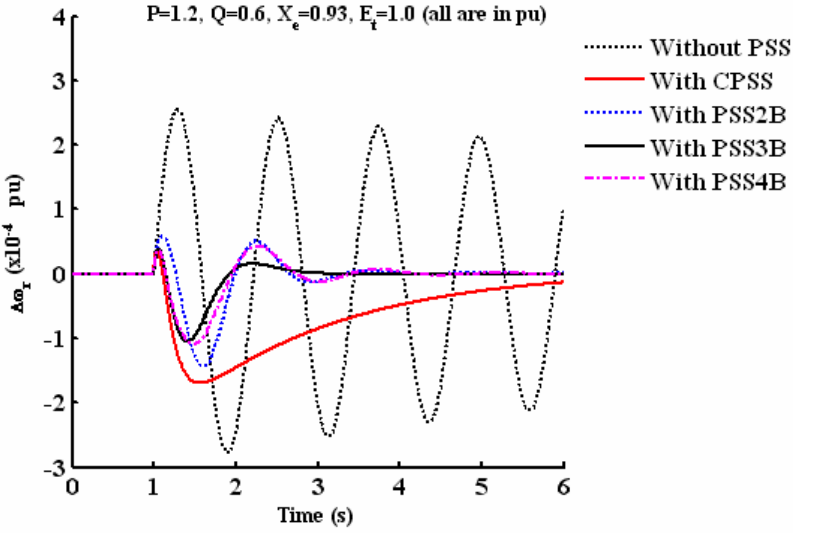


Fig. 27. Comparative responses of an SMIB model equipped with different PSS configurations and equipped with no PSS under 0.01 pu simultaneous step change in ΔV_{ref} and ΔT_m

9 Conclusion

In this study, only SMIB system is considered and drastic conditions considered are change in the operating conditions of active power flow, reactive power flow, change in LT bus voltage and change in line reactance and/or LT bus fault and subsequent opening of one of the two transmission lines. The IEEE dual input PSSs and classical phase compensation PSS are tested by evolutionary optimization techniques. From the study of the transient characteristics of lesser overshoot, lesser undershoot, lesser settling time, lesser size of state matrix, better damping and more robustness, it is established that the dual input IEEE-PSS3B produces the best optimal phase compensation as compared with classical conventional one. The feedbacks of electromagnetic torque deviation and angular speed deviation and processed through tunable lead-lag networks produce proper phase compensation and damping torque so that transient response becomes sufficiently damped for IEEE-PSS3B.

CASO and BFO are the two novel algorithms utilized for parameter tuning of the PSSs under both nominal and changing off-nominal system operating conditions. Any nominal controller based on any optimization technique is not necessarily fast enough for other drastically changing system operating conditions in real time. So, for scheduling of PSSs' parameters under off-nominal operating conditions, very fast, adaptive SFL is applied, SFL has very low computational burden.

List of Symbols

e_d	Field voltage
$E^{'}$	Voltage behind $X_d^{'}$
E_B	Infinite bus voltage
E_t	Generator's LT side bus voltage
fd	Field winding
H	Inertia constant, MW-s/MVA
K_D	Damping torque coefficient, pu torque/pu speed deviation
K_S	Synchronizing torque coefficient, pu torque/rad
o_{sh}	Overshoot of change in rotor speed, pu
$OF_1(\,)$	First objective function
$\mathrm{OF}_2(\,)$	Second objective function
P	Active power, pu
P_e	Electrical power output of the machine, pu
P_m	Mechanical power input to the machine, pu
Q	Reactive power, pu
R_T	Total system resistance, pu
t_{ex}	Execution time, s
t_{st}	Settling time of change in rotor speed, s
T_0	System constant
u_{sh}	Undershoot of change in rotor speed, pu
X_e	Equivalent transmission line reactance, pu
X_{Td}	Direct axis total reactance of the system, pu
X_{Tq}	Quadrature axis total reactance of the system, pu
$\Delta\delta$	Rotor angle deviation, elect rad

$\Delta\omega_r$	Speed deviation in pu = $\dfrac{\omega_r - \omega_0}{\omega_0}$
ΔT_e	Incremental change in electromagnetic torque, pu
ΔT_m	Incremental change in mechanical torque, pu
Δv_1	Incremental change in transducer output voltage, pu
ΔV_{pss}	Incremental change in PSS output voltage, pu
$\Delta V_{pss}^{\max}$	Maximum value of incremental change in PSS output voltage, pu
$\Delta V_{pss}^{\min}$	Minimum value of incremental change in PSS output voltage, pu
λ	Eigenvalue
ω_d , ω_n	Damped and undamped frequencies, respectively, rad/s
ω_0	Rated speed in elect rad/s = $2\pi f_0$
ψ_{fd}	Field flux linkage
ξ	Damping ratio
$\dfrac{d}{dt}(\Delta\omega_r)$	Time derivative of change in rotor speed

References

[1] Hariri, A., Malik, O.P.: A fuzzy logic based power system stabilizer with learning ability. IEEE Transaction on Energy Conversion 11(4), 721–727 (1996)

[2] Grigsby, L.L.: Electric power engineering handbook: power system stability and control. CRC Press, London (2007)

[3] Padiyar, K.R.: Analysis of subsynchronous resonance in power systems. Kluwer Academic Publishers, Boston (1999)

[4] Pal, Bikash, Chaudhuri, Balarko: Robust control in power systems. Springer, USA (2005)

[5] Padiyar, K.R.: Power system dynamics stability and control. BS Publications, India (2006)

[6] Kundur, P.: Power system stability and control. TMH, India (2006)

[7] El-Zonkoly, A.M.: Optimal tuning of power system stabilizers and AVR gains using particle swarm optimization. Expert Systems with Applications 31(3), 551–557 (2006)

[8] IEEE digital excitation task force of the equipment working group of the IEEE/PES excitation system subcommittee: Computer models for representation of digital-based excitation systems. IEEE Transaction on Energy Conversion 11(3), 607–615 (September 1996)

[9] Sauer, P.W., Pai, M.A.: Power system dynamics and stability. Pearson Education, India (2003)

[10] Radman, G., Smailli, Y.: Performance evaluation of PID power system stabilizer for synchronous generator. In: IEEE Conference, Southeastern (1988)
[11] Hopkins, W.E., Medanic, J., Perkins, W.R.: Output feedback pole placement in the design of suboptimal linear quadratic regulators. International Journal of Control 34(3), 593–612 (1981)
[12] Choic, S.S., Lim, M.: Design of wide range power system stabilizers via pole placement technique. Computers and Electrical Engineering 9(2), 103–110 (1982)
[13] Hsu, Yuan-Yin, Chen, Chao-Rong: Tuning of power system stabilizers using an artificial neural network. IEEE Transaction on Energy Conversion 6(4), 612–619 (1991)
[14] Soos, A., Malik, O.P.: An H2 optimal adaptive power system stabilizer. IEEE Transaction on Energy Conversion 17(1), 143–149 (2002)
[15] Hosseinzadeh, N., Kalam, A.: A direct adaptive fuzzy power system stabilizer. IEEE Transaction on Energy Conversion 14(4), 1564–1571 (1999)
[16] Abido, M.A.: Parameter optimization of multi-machine power system stabilizers using genetic local search. International Journal Electrical Power and Energy Systems 23(8), 785–794 (2001)
[17] Abido, M.A.: Robust design of multi-machine power system stabilizers using simulated annealing. IEEE Transaction on Energy Conversion 15(3), 297–304 (2000)
[18] Huang, T.L., Chang, C.H., Lee, J.L., Wang, H.M.: Design of sliding mode power system stabilizer via genetic algorithm. In: IEEE International Symposium Computational Intelligence in Robotics and Automation (2003)
[19] Goldberg, D.E.: Genetic algorithms in search, optimization and machine learning. Pearson Education, India (1989)
[20] Ghoshal, S.P., Goswami, S.K.: Application of GA based optimal integral gains in fuzzy based active power frequency control of non-reheat and reheat thermal generating systems. Electric Power Systems Research 67(2), 79–88 (2003)
[21] Ghoshal, S.P.: Application of GA/GA-SA based fuzzy automatic generation control of a multi-area thermal generating system. Electric Power Systems Research 70(2), 115–128 (2004)
[22] Passino, K.M.: Biomimicry of bacteria foraging for distributed automation and control. IEEE Control System Magazine 22(3), 52–67 (2002)
[23] Cai, Jiejin, et al.: Chaotic ant swarm optimization to economic dispatch. Electric Power Systems Research 77(10), 1373–1380 (2007)
[24] Mukherjee, V., Ghoshal, S.P.: Intelligent particle swarm optimized fuzzy PID controller for AVR system. Electric Power Systems Research 77(12), 1689–1698 (2007)
[25] Mukherjee, V., Ghoshal, S.P.: Comparison of intelligent fuzzy based AGC coordinated PID controlled and PSS controlled AVR system. International Journal Electrical Power and Energy Systems 29(9), 679–689 (2007)

A Metaheuristic Approach for Transmission System Expansion Planning

Ashu Verma, P.R. Bijwe, and B.K. Panigrahi

Abstract. Transmission network expansion planning (TNEP) is a very important problem to power system. The problem is very complex and computationally demanding because of discrete nature of decision variables and large number of options to be analyzed. Hence, more robust and efficient techniques are required to solve such problem. One of the recently developed metaheuristic, known as Harmony search (HS) has been tested for many engineering optimization problems and found to be an efficient tool for optimization. Hence, the motivation for this chapter is to investigate the application of HS for TNEP. The results for three sample test system are obtained to show the potential of the tested algorithm. The results obtained with HS are compared with basic binary GA to confirm the effectiveness of HS in terms of solution quality and number of fitness function evaluations required.

1 Introduction

An electric power system can be broadly divided in to three parts a) generating system, b) transmission system, and c) distribution system. The transmission system is responsible for transmitting electric power from generating station to distribution system or consumer level. The demands and generations in a power system are increasing day by day. This growth of electricity gives rise to the need of transmission system planning. The objectives of transmission network expansion planning (TNEP) are based on existing systems, future load, generation scenarios, available right- of-ways, cost of line etc. The TNEP is an important part of power system planning. It consists of determining the optimal expansion plan of transmission network such that the total cost of new constructed transmission lines is minimum while satisfying the operational constraints of the power system. The TNEP is a large scale, non linear, non convex, mixed integer optimization problem. It has large number of options to be analyzed. Further, the number of options to be analyzed increases exponentially with the size of the system. Hence, large number of sub optimal solutions exist for this problem. TNEP can be divided in to two types 1) static TNEP and 2) dynamic TNEP. Static TNEP performs all the expansions in a single stage of planning horizon. However, Dynamic TNEP decides when, where and how many new circuits should be installed to serve the growing electricity demand in an optimal way. The static TNEP being already very diffi-

Ashu Verma, P.R. Bijwe, and B.K. Panigrahi
Department of Electrical Engineering,
Indian Institute of technology, Delhi

B.K. Panigrahi et al. (Eds.): Computational Intelligence in Power Engineering, SCI 302, pp. 367 – 379.
springerlink.com

cult, the dynamic TNEP becomes even more complex. Secondly, the restructuring of electricity market has added even more to the complexity of TNEP problem.

The metaheuristic techniques prove to be very well suited for such type of combinatorial optimization problems. Many techniques such as genetic algorithm (GA), particle swarm optimization (PSO), simulated annealing (SA), tabu search (TS) has been proposed and tested for transmission network expansion planning. HS being a recently developed metaheuristic, has been tested for variety of engineering optimization problem and found to be very successful. However, its investigation for TNEP has not been done so, far. Hence, the motivation in this chapter is to investigate the usefulness of HS algorithm.

The HS algorithm for solving engineering optimization problems was first presented in (Geem et.al. 2001). The algorithm is based on the musical process of searching for the perfect state of harmony. The harmony in the music is analogous to the optimization solution vector and improvisations are analogous to the global and local search schemes. Compared to earlier meta-heuristic, it imposes fewer mathematical requirements and can be easily adopted for various engineering optimization problems. Application of HS for some of the problem can be found in (Kim et.al. 2001, Geem 2009, kang et.al. 2004, Geem et.al. 2005). An improved version of HS has been found in (Mahdavi et.al. 2007), where it employs a novel method for generating new solution vector that enhances the accuracy and convergence of the classical HS.

The improved HS (IHS) presented in (Mahdavi et.al. 2007) is used in this chapter to demonstrate its application in TNEP problem.

2 Transmission Network Expansion Planning

In this chapter only static TNEP has been investigated. Since, TNEP problem is too complex and difficult to be fully dealt in a single chapter. Hence, a simple DC model is used for static TNEP where, the cost of transmission lines to be constructed is the only objective and real power flows in the lines in base case and *N-1* contingency cases are the only constraints. The *N-1* contingency analysis looks at the system state after a single line outage and is used to access the system security. However, the algorithm is general enough to consider other aspects of transmission expansion planning like operational economy, deregulation, dynamic TNEP etc.

The basic model for static TNEP with security constraints explained in (Silva et.al. 2005) is used as a base for formulating TNEP in this chapter. The TNEP with security constraints can be represented as

$$\min \; v = \sum_{l \in \Omega} c_l n_l \tag{1}$$

$$S \; f^k + g \; = d \tag{2}$$

$$f_l^k - \gamma_l \left(n_l^0 + n_l \right) (\Delta \theta_l^k) = 0, \text{for } l \in 1,2....., nl \;\&\; l \neq k \tag{3}$$

$$f_l^k - \gamma_l \left(n_l^0 + n_l - 1 \right) (\Delta \theta_l^k) = 0, \text{for } l{=}k, \tag{4}$$

$$\left|f_l^k\right| \le \left(n_l^0 + n_l\right)\overline{f_l} \quad \text{,for } l \in 1,2\text{.....}, nl \,\&\, l \ne k \tag{5}$$

$$\left|f_l^k\right| \le \left(n_l^0 + n_l - 1\right)\overline{f_l}\,, \quad \text{for } l=k, \tag{6}$$

$$0 \le n_l \le \overline{n_l} \tag{7}$$

$n_l \ge 0,$ and integer, for $l \in 1,2\text{.....}, nl \,\&\, l \ne k$

$(n_l + n_l^0 - 1) \ge 0,$ and integer, for $l = k$

$l \in \Omega$ and k=0,1.....*NC*

where, k=0, represnts the base case without any line outage.

S : branch-node incidence transposed matrix of the power system,

f^k :vector with elements f_l^k,

γ_l : susceptance of the circuit that can be added to l^{th} right-of-way,

n_l : the number of circuits added in l^{th} right-of-way,

n_l^0 : no. of circuits in the base case,

$\Delta\theta_l^k$: phase angle difference in lth right-of way when kth line is out,

f_l^k : total real power flow by the circuit in lth right-of-way when kth line is out,

$\overline{f_l}$: maximum allowed real power flow in the circuit in lth right-of-way ,

$\overline{n_l}$:maximum number of circuits that can be added in lth right-of-way,

Ω : set of all right of ways,

nl : total number of lines in the circuit,

NC :number of credible contingencies(taken as equal to *nl* in the present case).

The objective of TNEP problem is to minimize the total investment cost of the new transmission lines to be constructed, satisfying the constraint on real power flow in the lines of the network, for base case and *N-1* contingency cases. Constraint (2) represents the power balance at each node. Constraint (3) and (4) are the real power flow equations in DC network. Constraint (5) and (6) represents the line real power flow constraint. Constraint (7) represents the restriction on the construction of lines per corridor (R.O.W).

3 Improved Harmony Search Algorithm (IHS)

The harmony search algorithm is based on an analogy with music improvisation process, where music players improvise the pitches of their instruments to obtain better harmony (Geem et.al. 2001). The musical performers try to find the perfect state of harmony as determined by aesthetic standard, just as the optimization process try to find a global optimum solution. The pitch of each musical instrument determines the aesthetic quality, just as the objective function value is

determined by the set of values assigned to each decision variable. The IHS algorithm proposed in [5] is used for solving the STNEP problem in this chapter. Brief outline of the method is as follows.

The following steps are involved in any HS algorithm

1. Initialize the problem and algorithm parameters.
2. Initialize the harmony memory.
3. Improvise a new harmony.
4. Update the harmony memory.
5. Check for the stopping criteria.

For better understanding of the algorithm, these steps are briefly described in the following five subsections.

A. Initialize the problem and algorithm parameters.
In this step, we specify the optimization problem as follows:

$$\text{Minimize } f(x) \tag{8}$$
$$\text{Subject to } x_i \in X_i \ , \ i=1,2.....N \tag{9}$$

Where, $f(x)$ is an objective function, x is set of decision variables x_i, N is the number of decision variables, X_i represents the possible range of values for each decision variables.

The HS algorithm parameters to be initialized are as follows:

1. Harmony memory size (HMS) which indicates the number of solution vectors in the harmony memory.
2. Harmony memory considering rate (HMCR):
3. Pitch adjustment rate (PAR).
4. Number of improvisations (NI) or stopping criteria.

B. Initialize the harmony memory.
The harmony memory is initialized with as many randomly generated vectors as the value of HMS.

$$\text{HM} = \begin{bmatrix} x_1^1 & x_2^1 & \ldots\ldots & x_{N-1}^1 & x_N^1 \\ x_1^2 & x_2^2 & \ldots\ldots & x_{N-1}^2 & x_N^2 \\ \vdots & \vdots & \ldots\ldots & \vdots & \vdots \\ x_1^{\text{HMS-1}} & x_2^{\text{HMS-1}} & \ldots\ldots & x_{N-1}^{\text{HMS-1}} & x_N^{\text{HMS-1}} \\ x_1^{\text{HMS}} & x_2^{\text{HMS}} & \ldots\ldots & x_{N-1}^{\text{HMS}} & x_N^{\text{HMS}} \end{bmatrix} \tag{10}$$

C. Improvise a new harmony.
The new harmony vector $x^t = (x_1^t, x_2^t .....x_i^t)$ is generated based of three rules: (1) memory consideration, (2) pitch adjustment and (3) random selection.

$$x_i^t \leftarrow \begin{cases} x_i^t \in \left\{ x_i^1, x_i^2 .....x_i^{HMS} \right\} \text{with probability HMCR,} \\ x_i^t \in X_i \text{ with probability (1-HMCR)} \end{cases} \tag{11}$$

A HMCR of 0.92 indicates that the HS algorithm will choose the decision variable from the stored values in the HM with 92% probability and from the entire range with ((100-92)%) probability. Every component chosen by harmony consideration is examined for pitch adjustment based on the following rule.

Pitch adjusting decision for x_i^t is given as

$$x_i^t \leftarrow \begin{cases} \text{Yes with probability PAR,} \\ \text{No with probability (1-PAR)} \end{cases} \tag{12}$$

The value of (1-PAR) defines the rate of doing nothing. If pitch adjustment decision for x_i^t is Yes, x_i^t is modified as follows

$$x_i^t = x_i^t \pm \text{rand}(\,) * \text{bw}\,, \tag{13}$$

Where, bw is an arbitrary distance bandwidth, rand( ) is a random number between 0 and 1.

The values of PAR are adjusted as follows,

$$\text{PAR(gn)} = \text{PAR}_{\min} + \frac{(\text{PAR}_{\max} - \text{PAR}_{\min})}{NI} X \text{ gn} \tag{14}$$

Where,
gn: an improvisation or generation
gn=1,2……*NI*,
PAR(gn); pitch adjusting rate for generation or improvisation gn,
$\text{PAR}_{\min}$, $\text{PAR}_{\max}$: minimum and maximum pitch adjusting rate,
And the bw is calculated as follows,
$\text{bw(gn)}=\text{bw}_{\max}\exp(\text{c.gn})$

$$\text{c}=\frac{\text{Ln}\left(\frac{\text{bw}_{\min}}{\text{bw}_{\max}}\right)}{\text{NI}} \tag{15}$$

bw(gn): bandwidth for each generation,
$\text{bw}_{\min}$: minimum bandwidth,
$\text{bw}_{\max}$: maximum bandwidth.

D. Update the harmony memory.

The new memory is assessed in terms of the objective function (fitness function) value and if the new memory is better than the worst memory in the HM, the new harmony memory is included in the HM and the existing worst harmony is excluded from the HM.

E. Check for stopping criteria:

If maximum number of improvisations is reached, then stop, otherwise steps *C* and *D* are repeated.

HS for STNEP

The *f(x)* represents the objective function represented by equation (1) for STNEP. *x* defines the set of candidate lines presenting a solution to STNEP. Each element in *x* represents the right-of-way in which a candidate line is to be constructed. The range of each variable defined by X_i indicates the list of available right-of-ways. If two lines are added in a particular right-of-way, then two elements with same number (indicating same right-of-way) will come in the vector *x*.

Fitness function evaluation:

To check the worth of a vector (solution) in the harmony memory, fitness function is evaluated using the following equation.

$$f = \sum_{l} c_l n_l + \mathrm{W}_1 \sum_{k=0}^{NC} \sum_{ol} (\text{abs } (f_l^k) - \overline{f_l}) \; + \mathrm{W}_2 (n_l - \overline{n_l}) \tag{16}$$

ol: represents the set of overloaded lines.

The objective of the STNEP represented by equation (1)-(7) is to find the set of transmission lines to be constructed such that the cost of expansion plan is minimum and no overloads are produced during the planning horizon. Hence, first term in the equation (16) indicates the total cost of transmission lines to be constructed. The second term is added to the objective function for the real power flow constraint violations in the base case, and *N-1* contingency cases. The third term is added to the objective function if number of circuits added in l^{th} right-of-way exceeds the maximum limit. The second and third terms are added to the fitness function only in case of violations. W_1, W_2 are constants.

4 Results and Discussions

The proposed algorithm has been tested for three standard test systems, 6 bus Garver system, IEEE 24 bus system, and South Brazilian 46 bus system. The comparison of results is presented with the one obtained with basic binary GA to show the potential of the proposed approach. An algorithm for basic binary GA is implemented to compare the results.

4.1 Garver 6 Bus System

This system has six buses, 15 candidate branches, a total demand of 760 MW, and a maximum possible number of lines per corridor is equal to five. The initial topology and electrical system data is found in (Silva et.al. 2005). The initial topology is also shown in Fig. 1 for reference.

4.1.1 TNEP without Security Constraints

The optimal solution to the expansion planning problem without security constraints results in a total investment cost of US$200,000 with the addition of

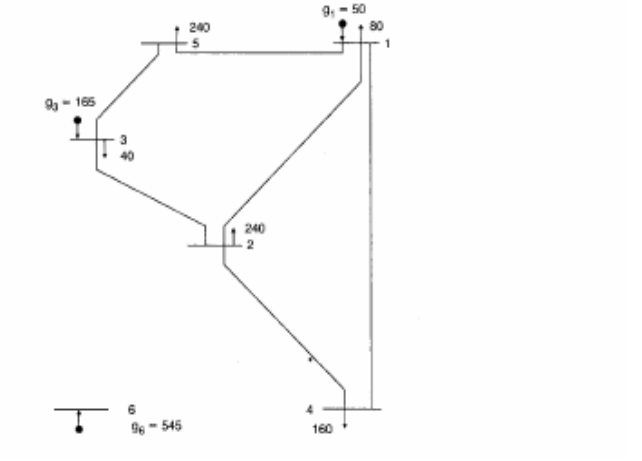


Fig. 1. Garver 6 bus system

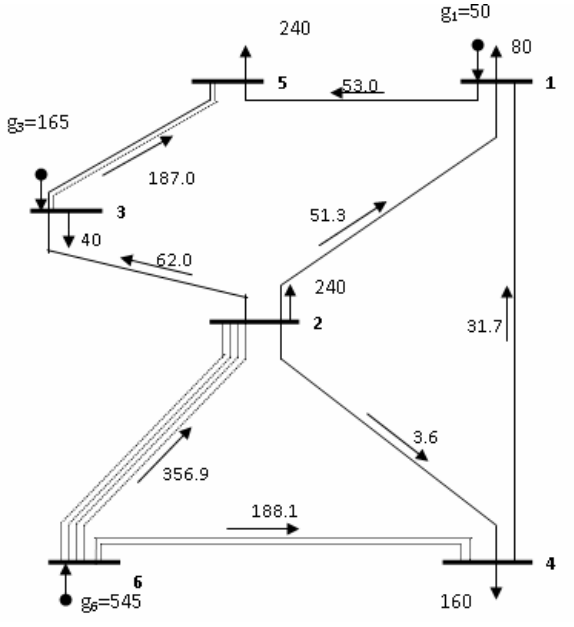


Fig. 2. Garver 6 bus system optimal expansion plan for TNEP without security constraints

following seven lines, $n_{2\text{-}6}$=4, $n_{4\text{-}6}$=2 and $n_{3\text{-}5}$=1. The final expansion plan is shown in Fig.2. The broken lines are the planned one. The results obtained with the proposed algorithm matches exactly with the results obtained with GA and the one reported in (Silva et.al. 2005).

4.1.2 TNEP with Security Constraints

The optimal expansion plan with security constraints results an investment cost of US$298 000, with the addition of total ten lines as follows, n_{2-6}=4, n_{3-5}=2, n_{3-6}=1, and n_{4-6}=3. The final expansion plan is shown in Fig. 3. The results obtained with the proposed algorithm matches exactly with the results obtained with GA and the one reported in (Silva et.al. 2005).

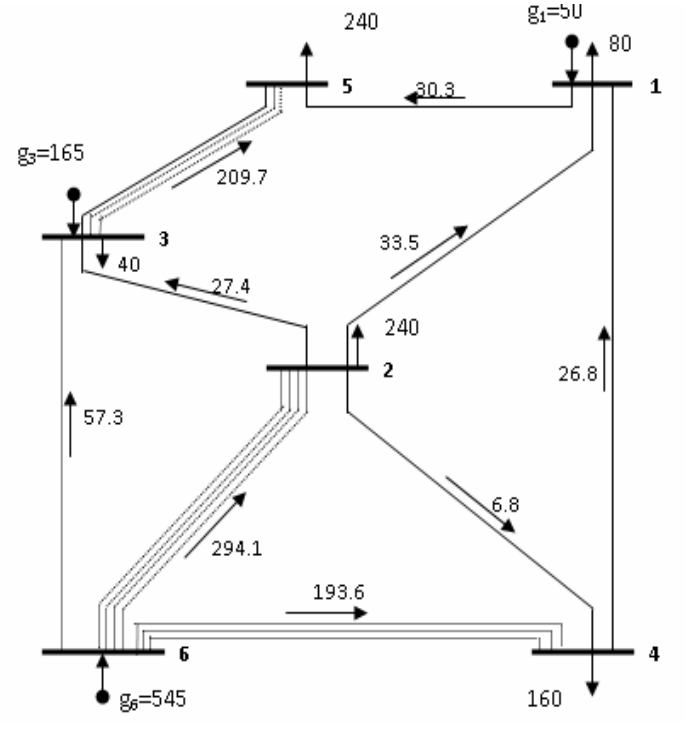


Fig. 3. Garver 6 bus system optimal expansion plan for TNEP with security constraints

4.2 IEEE 24 Bus System

This system consists of 24 buses, 41 candidate circuits and 8550 MW of total demand. The maximum number of lines allowed per corridor is four. The initial network and the electrical data can be found in (Fang et.al. 2003). The generation/load data has been taken for Plan G_3 of (Romero et.al. 2005). The initial network is also shown in Fig. 4 for reference.

4.2.1 TNEP without Security Constraints

The final optimal solution obtained with HS for TNEP without security constraints results in an investment cost of US$ 214×10^6 with the addition of seven lines as follows.

n_{16-19}=1, n_{14-16}=1, n_{6-10}=1, n_{7-8}=2, n_{10-12}=1, n_{20-23}=1.

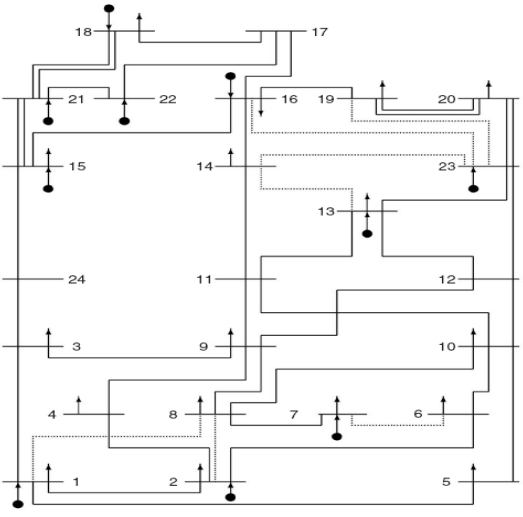


Fig. 4. IEEE 24 bus system

The cost of expansion plan obtained with the proposed method is 1.8 % lesser than the cost obtained from the plan presented in (Romero et.al. 2005) and matches exactly with the one obtained with GA.

4.2.2 TNEP with Security Constraints

The final optimal solution obtained with HS for TNEP with security constraints results in an investment cost of US \$869 x10^6 with the addition of twenty one lines as follows.

$n_{1\text{-}5}$=2, $n_{3\text{-}9}$=2, $n_{4\text{-}9}$=1, $n_{5\text{-}10}$=2, $n_{6\text{-}10}$=2, $n_{7\text{-}8}$=3, $n_{9\text{-}12}$=1, $n_{10\text{-}12}$=1, $n_{12\text{-}23}$=1, $n_{13\text{-}14}$=1, $n_{14\text{-}23}$=1, $n_{15\text{-}21}$=1, $n_{18\text{-}21}$=1, $n_{20\text{-}23}$=1, $n_{21\text{-}22}$=1.

The solution obtained with basic binary GA results in an investment cost of US \$903 x10^6 with the addition of following twenty lines,

$n_{7\text{-}8}$=3, $n_{6\text{-}10}$=3, $n_{1\text{-}5}$=2, $n_{15\text{-}21}$=1, $n_{20\text{-}23}$=1, $n_{3\text{-}24}$=1, $n_{16\text{-}19}$=1, $n_{14\text{-}23}$=2, $n_{15\text{-}24}$=1, $n_{9\text{-}12}$=1, $n_{10\text{-}12}$=1, $n_{12\text{-}13}$=1, $n_{19\text{-}22}$=1, $n_{4\text{-}9}$=1.

The parameters used for HS, and GA for this study are as follows

For HS: harmony memory size (HMS) =50 ; harmony memory considering rate (HMCR)=0.98; maximum pitch adjustment rate (PARmax)=0.99; minimum pitch adjustment rate (PARmin)=0.1; number of improvisations (NI) or stopping criteria=3000.

For GA: population size=750,crossover rate=0.8, mutation rate=0.01, number of generations=4000.

The number of fitness function evaluations required by both algorithms for the above study are given in Table 1.

Table 1. Comparison of cost and number of fitness functions required for optimal solution with two methods for IEEE 24 bus system

Method	No.of fitness function Evaluations	Cost of Expansion Plan [US$ 10^6]
HS	**1, 70,400**	**869**
Basic Binary GA	27,53,166	903

It can be observed from the results that HS provides much better results (lower cost) with far lesser number of fitness function evaluations as compared to those with basic binary GA. The cost of expansion plan obtained with HS is 3.77% lesser than that obtained with GA. The number of fitness function evaluations required by HS is 93.85 %, lesser as compared to that required by GA.

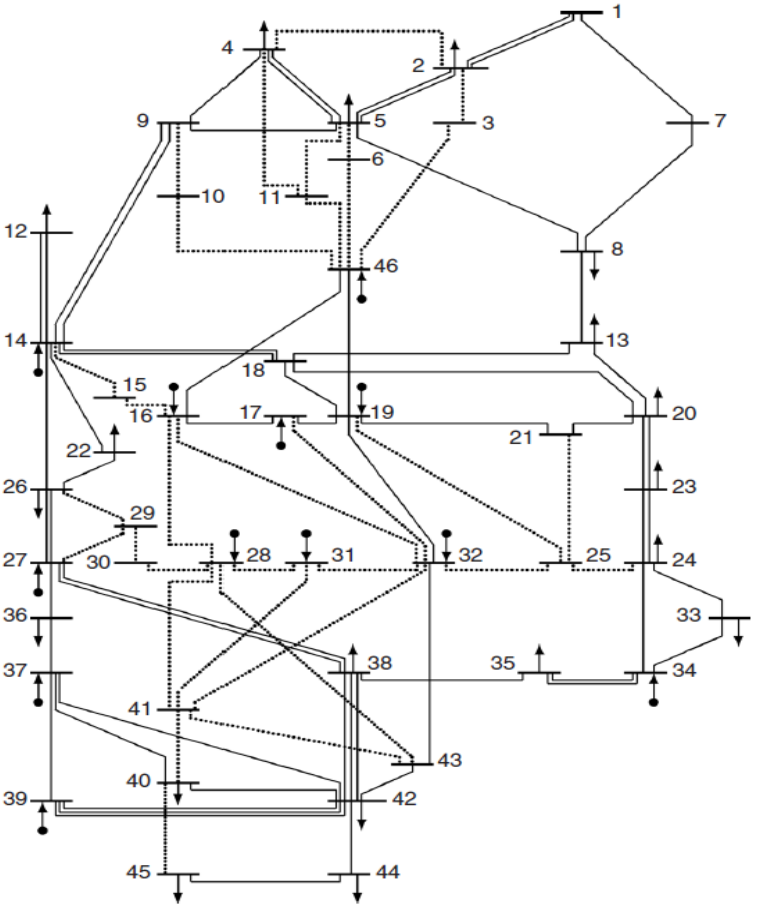


Fig. 5. Basic configuration of 46 bus South Brazilian System

4.3 South Brazilian 46 Bus System

This system has 46 buses, 79 right-of-ways on which circuits can be constructed, and a total demand of 6880 MW. The maximum number of lines which can be added to each corridor is six. This is a realistic system representing a good test case for the proposed algorithm. The relevant data can be found in (Haffner et.al. 2000). The initial network is taken from (Haffner et.al. 2000) and is also shown in Fig.5 for ready reference.

4.3.1 TNEP without Security Constraints

The final optimal solution obtained with HS for TNEP without security constraints results in an investment cost of US$154 420,000 which matches exactly with the one reported in (Romero et.al. 1994).

The lines present in the solution are as follows

$n_{20\text{-}21}$=1, $n_{42\text{-}43}$=2, $n_{31\text{-}32}$=1, $n_{28\text{-}30}$=1, $n_{26\text{-}29}$=3, $n_{29\text{-}30}$=2, $n_{24\text{-}25}$=2, $n_{19\text{-}25}$=1, $n_{46\text{-}6}$=1, $n_{6\text{-}5}$=2.

4.3.2 TNEP with Security Constraints

The final optimal solution obtained with HS for TNEP with security constraints results in an investment cost of US $343,870,000 with the addition of following thirty lines.

$n_{26\text{-}29}$=3, $n_{29\text{-}30}$=2, $n_{28\text{-}30}$=1, $n_{31\text{-}32}$=1, $n_{24\text{-}25}$=4, $n_{25\text{-}32}$=1, $n_{25\text{-}21}$=1, $n_{46\text{-}6}$=2, $n_{6\text{-}5}$=4, $n_{2\text{-}5}$=1, $n_{12\text{-}14}$=1, $n_{17\text{-}19}$=1, $n_{19\text{-}21}$=1, $n_{20\text{-}21}$=2, $n_{26\text{-}27}$=1, $n_{27\text{-}38}$=1, $n_{32\text{-}43}$=1, $n_{42\text{-}43}$=3.

The final optimal solution obtained with GA results in an investment cost of US$ 432 350,000 with the addition of following forty three lines,

$n_{26\text{-}29}$=3, $n_{29\text{-}30}$=2, $n_{28\text{-}30}$=1, $n_{31\text{-}32}$=1, $n_{46\text{-}11}$=1, $n_{11\text{-}5}$=3, $n_{24\text{-}25}$=4, $n_{25\text{-}32}$=1, $n_{25\text{-}21}$=1, $n_{46\text{-}6}$=2, $n_{6\text{-}5}$=2, $n_{12\text{-}14}$=3, $n_{2\text{-}5}$=2, $n_{16\text{-}28}$=1, $n_{26\text{-}27}$=1, $n_{1\text{-}7}$=2, $n_{19\text{-}21}$=1, $n_{9\text{-}14}$=1, $n_{27\text{-}36}$=1, $n_{28\text{-}43}$=1, $n_{42\text{-}43}$=3, $n_{37\text{-}39}$=2, $n_{34\text{-}35}$=1, $n_{20\text{-}21}$=2, $n_{2\text{-}4}$=1.

A rigorous study of algorithm parameters done by varying them for the above system between the permissible ranges is given in Table 2 and 3. Initially the effect of variations of HMCR is observed by keeping HMS constant. Then with the best value of HMCR obtained, the effect of variations of HMS is observed.

It can be observed from Table 2 and 3 that HMCR=0.98 and HMS =50 provided best results in terms of no of fitness function evaluations and standard deviation for the above system. However, these settings are system dependent. With experiments in large number of systems with different levels of complexity, it has been observed that for TNEP problem, HMCR value of 0.98 provides best results. However, HMS needs to be adjusted depending upon the size of the system and level of complexity of TNEP problem.

Hence, the algorithm parameters used for this system are as follows: harmony memory size (HMS) =50; harmony memory considering rate (HMCR) =0.98; maximum pitch adjustment rate (PARmax)=0.99; minimum pitch adjustment rate (PARmin)=0.1; number of improvisations (NI) or stopping criteria=3000.

Table 2. Effect of HMCR variation on performance of HS algorithm for TNEP (with HMS=50)

HMCR	**0.99**	**0.98**	**0.95**	**0.9**
Cost of expansion Plan [US$ 10^6]	343. 870	343. 870	343. 870	349.17
Standard Deviation	31.39	**25**	18.85	54
No. of fitness function evaluations	439750	**296800**	354600	318700

Table 3. Effect of HMS variation on performance of HS algorithm for TNEP (with HMCR=0.98)

HMS	**25**	**50**	**75**	**100**
Cost of expansion Plan [US$ 10^6]	343. 870	343. 870	343. 870	343. 870
Standard Deviation	50	**25**	25.40	30.54
No. of fitness function evaluations	140500	**296800**	321575	482500

For GA. no. of population=1000, crossover rate=0.8, mutation rate=0.01, number of generations=4000.

The number of fitness function evaluations required by both methods for the results shown above are given in Table 4.

Table 4. Comparison of cost and number of fitness functions required for optimal solution with different methods for 46 bus South Brazilian system

Method	No.of fitness function evaluations	Cost of Expansion Plan [US$ 10^6]
HS	**2.96×10^5**	343. 870
Basic Binary GA	2.67×10^6	432.350

It can be observed from the above results that the HS algorithm provides much better (low cost) results with lesser number of fitness function evaluations as compared to basic binary GA. The cost of expansion plan obtained with IHS is 20.46% lesser than the one obtained with GA. The number of fitness function evaluations required by HS is 8.89 % lesser as compared to that required by GA.

5 Summary and Conclusions

It can be found from the literature that HS has been proved to be better algorithm for handling complex engineering optimization problems. However, its usefulness for TNEP has not been investigated so far. Hence, in this chapter the application of HS for TNEP has been presented. The comparison of results is carried out with the one obtained with basic binary GA. Results for three sample test systems confirm the potential of the tested algorithm. The discussed algorithm provides better (low cost) solution in all the cases with lesser number of fitness function evaluations.

References

Geem, Z.W., Kim, J.H., Loganathan, G.V.: A new heuristic optimization algorithm: harmony search. Simulation 76(2), 60–68 (2001)

Kim, J.H., Geem, Z.W., Kim, E.S.: Parameter estimation of the nonlinear Muskingum model using harmony search. J. Am. Water Resour. Assoc. 37(5), 1131–1138 (2001)

Geem, Z.W.: Harmony search optimization to Pump-Included Water Distribution Network design. Civil Engineering and Environmental systems 26(3), 211–221 (2009)

Kang, S.L., Geem, Z.W.: A new structural optimization method based on the harmony search algorithm. Comput. Struct. 82(9-10), 781–798 (2004)

Geem, Z.W., Tseng, C., Park, Y.: Harmony search for generalized orienteering problem: best touring in China. LNCS, vol. 3412, pp. 741–750. Springer, Heidelberg (2005)

Mahdavi, M., Fesanghary, M., Damangir, E.: An improved harmony search algorithm for solving optimization problems. Applied Mathematics and Computation 188, 1567–1579 (2007)

de Silva, I.J., Rider, M.J., Romero, R., Garcia, A.V., Murari, C.A.: Transmission network expansion planning with security constraints. IEE Proc. Gener. Transm. Distrib. 152(6), 828–836 (2005)

Fang, R., Hill, D.J.: A new strategy for transmission expansion in competitive electricity markets. IEEE Trans. Power Syst. 18(1), 374–380 (2003)

Romero, R., Rocha, C., Mantovani, J.R.S., Sánchez, I.G.: Constructive heuristic algorithm for the DC model in network transmission expansion planning. IEE Proc.-Gener. Transm. Distrib. 152(2), 277–282 (2005)

Haffner, S., Monticelli, A., Garcia, A., Mantovani, J., Romero, R.: Branch and bound algorithm for transmission system expansion planning using a transportation model. IEE Proc. Gener. Transm. Distrib. 147(3), 149–156 (2000)

Romero, R., Monticelli, A.: A hierarchical decomposition approach for transmission network expansion planning. IEEE Trans. Power Syst. 9(1), 373–380 (1994)

Author Index