FAULT LOCATION ESTIMATION BASED ON MATCHING THE SIMULATED AND RECORDED WAVEFORMS USING GENETIC ALGORITHMS

M Kezunovic, Y Liao *

Texas A&M University, College Station, USA

I. INTRODUCTION

Prompt and accurate location of the faults in a large-scale transmission system is critical when system reliability is considered and usually is the first step in the system restoration. The accuracy of fault location estimation essentially depends on the information available. While there have been some successful algorithms for fault location utilizing two-end or three-end data, the satisfactory solutions are harder to formulate if only the local information or only the data at limited substation locations are available [Girgis et al (1), Waikar et al (2)].

To improve the accuracy for fault location when only limited recorded data are available, the "waveform matching" based approach may be used. In this approach, simulation studies are carried out to obtain simulated waveforms under specified fault conditions. The simulated waveforms are then compared with the recorded ones. By iteratively posing faults in the system, running simulations, and comparing the simulated waveforms with the recorded ones, an optimal estimate of the fault location may be obtained. It may be determined as the one specified in the simulation studies that allows simulating the waveforms that best match the recorded ones. The matching is made at the phasor level presently.

In this paper, the fault location estimation is mathematically formulated as an optimization problem of which the fault location and fault resistances are unknown variables. An efficient GA based searching scheme is developed for obtaining the solution that is globally optimal [Goldberg (3)].

The rest of the paper is organized as follows. The problem statement is presented first. Key concepts, namely "sparse data" and "waveform matching" are illustrated. The proposed new genetic algorithm based implementation approach is illustrated next. Finally, evaluation studies are carried out to verify the accuracy and feasibility of the proposed approach.

II. PROBLEM STATEMENT

The proposed approach will make use of the "waveform matching" based methodology. The two key concepts, namely the "sparse data" and "waveform matching" are illustrated as follows.

A. Sparse data

Sparse data, in our work, is referred to the data obtained from recording devices sparsely located at various substation locations. Examples of recording devices may include digital fault recorders (DFR), digital relays, or other intelligent electronic devices (IED) [Kezunovic (4)]. The data captured by recording devices may include analog quantities such as voltage and current waveforms and digital quantities such as breaker status and relay operation status. Both analog and digital quantities may be useful for locating the fault [4].

If only sparse data are available for fault location, in many cases, none of the one-end, two-end and three-end algorithms may be applicable for locating the faults with satisfactory accuracy [1-2, 4]. To solve the fault location problem utilizing sparse data, the "waveform matching" based approach may be used as illustrated next.

B. Waveform matching

For improved accuracy of fault location by utilizing sparse data, the waveform-matching based approach may be utilized. The model of the power system is utilized to carry out simulation studies. The matching is made between the voltage and current waveforms obtained by recording devices and those generated in simulation studies. The fault is searched through the system by utilizing an iterative searching process. The searching process may consist of the following steps. First an initial fault location is assumed. Second, the simulation studies are set up according to the specified fault. Third, the simulation studies corresponding to the specified fault are carried out utilizing appropriate simulation tools. Fourth, simulated waveforms of quantities of interest are obtained. Fifth, the simulated waveforms are compared with recorded ones, and the matching degree of the simulated and recorded waveforms is evaluated by using appropriate criteria. Sixth, the initial fault location is modified according to certain approaches, and then the process proceeds to the second step and continues. The above steps are iterated until the simulated waveforms that best match the recorded ones are produced. The fault location will be determined as the one specified in the simulation studies when generating simulated waveforms that best match the recorded ones.

To evaluate the matching degree of the simulated and recorded waveforms, two different criteria may be employed. The first one utilizes phasors for matching.

* Y Liao is presently employed by ABB-ETI

The other one utilizes transients for matching. Only the phasor matching is investigated in this paper. The short circuit studies may be carried out to obtain the phasors under the specified fault conditions. Then the phasors are compared with those derived from the recorded waveforms.

To run simulation studies, short-circuit model of the system is needed. Among other existing software packages, Power System Simulator for Engineering (PSS/E) may be used for carrying out short circuit studies [PTI (5)].

So far, the "waveform matching" based approach for fault location has already been employed manually by some engineers. The matching is made at the phasor level. There may be certain difficulties involved in the manual-matching based methods. First, it may be time-consuming and even difficult to manually pose faults, run simulations and compare the simulated and recorded phasors. The tedious process may be prone to human errors. Second, there is no accepted approach for guiding the searching process. The engineers usually have difficulty in knowing where to pose faults in the next iterative step, and may have to pose faults randomly. Third, since the fault resistance is unknown, a zero fault resistance usually has to be assumed in manual methods when posing faults. This may introduce undesirable errors. Due to the limitations of manual matching, the "waveform matching" based approach has not been widely used in practice. It has not received wide research attention either. There has not been a systematic approach available in the literature for automatically implementing the concept. Our research aims at proposing a new implementation method for automated "phasor matching" that may facilitate the entire searching process and lead to improved accuracy.

III. PROPOSED IMPLEMENTATION APPROACH FOR PHASOR MATCHING

To effectively guide the searching process, a genetic algorithm (GA) based searching approach based on "phasor matching" is proposed. In the following sections, the fault location is first formulated as an optimization problem, and then the application of GA for fault location estimation is illustrated.

A. Formulation of the fault location searching process as an optimization problem

In the fault location problem discussed so far, there may be two possible unknown parameters, namely the fault location and the fault resistance. In our work, the type of the fault is assumed to be obtained by another fault type classification program.

To search for the fault location by phasor matching, the two unknown parameters, i.e. fault location and resistance, need to be varied. By specifying faults with various fault resistances and locations, a number of short-circuit studies can be carried out and the corresponding voltage and current phasors obtained. The most probable solution to the problem may be determined as the one specified in the simulation studies when producing the phasors that best match those derived from the recorded waveforms.

The formulation of the problem can be presented as follows:

Find the value of x and R_f that minimize

$$f_c(x, R_f) = \sum_{k=1}^{N_v} \{|V_{ks} - V_{kr}|\} + \sum_{k=1}^{N_i} \{|I_{ks} - I_{kr}|\} \qquad (1)$$

or maximize

$$f_f(x, R_f) = -f_c(x, R_f) \qquad (2)$$

where

$f_c(x, R_f)$: the defined cost function

$f_f(x, R_f)$: the defined fitness function. The larger the value of the fitness function, the better the matching of the simulated and recorded phasors, and the better the solution is.

x : the fault location

R_f : fault resistance

V_{ks} and V_{kr} : the during-fault voltage phasors obtained from short-circuit simulation studies and from recorded waveforms respectively.

I_{ks} and I_{kr} : the during-fault current phasors obtained from short-circuit simulation studies and from recorded waveforms respectively.

k : the index of the voltage or current phasors

N_v and N_i : the total number of the voltage and current phasors respectively.

It is noted that the largest possible fitness value defined by (2) is equal to zero and can be reached if the phasors obtained from simulation studies exactly match those obtained from recorded waveforms. Therefore, the best fault location estimate would be the one that maximizes (2). Appropriate optimization techniques need to be selected to solve this problem.

To obtain a clear picture of the nature of the fitness function, various simulation studies have been carried out to obtain the fitness value versus the fault location and resistance using the sample power system shown in Fig. 1. For convenience of presentation, all the data are assumed to be obtained by DFRs in this work [4]. The depicted system represents a portion of the 138 kV Reliant Energy HL&P transmission system. In our work, the commercial software package PSS/E is utilized to carry out the short circuit studies [5]. The fitness value is obtained by specifying the faults with various fault resistance on each line throughout the system, running simulations, and applying (1-2).

Now assume that a phase-A to ground fault with fault resistance 0.1 p.u. occurs on the line between bus 13 and 12 that is 9.1 miles away from bus 13. If only the recorded data at bus 1 are available for fault location, the fitness value versus the fault location and fault resistance for this specific fault can be obtained as depicted in Fig. 2.

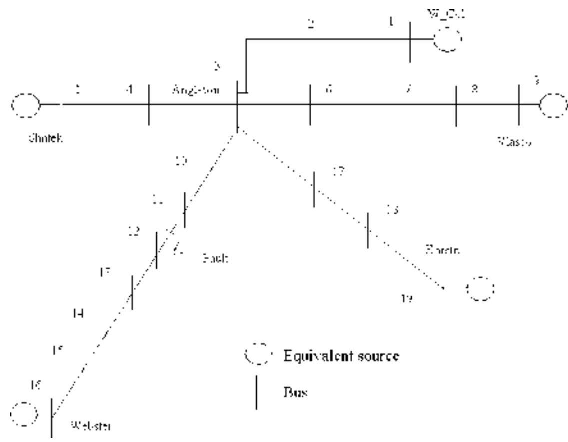

Figure 1: A sample power system

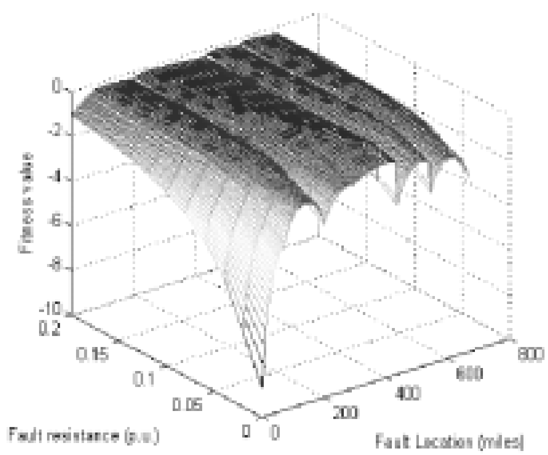

Figure 2: The fitness surface for an a-g fault using data at bus 1 and (2)

It is seen from Fig. 2 that the maximum fitness value occurs at point (312.7, 0.1), which is the optimal solution for the phase A to ground fault.

It is noted that the fitness surface is not regular and contains saddle points and local maximum points. Simulation studies evince similar characteristics for other types of faults. Hence, it is rather difficult to use the gradient-based method to find the global maximum point. Exhaustive search through every possible solution may be too time-consuming and hence impractical.

The GA based optimization approach is good at finding the globally optimal solution and avoiding the local optima. The nature of the fitness function, as depicted in Fig. 2, prompts the attempt of the GA based optimization method as described next [3].

B. Proposed genetic algorithm based searching approach

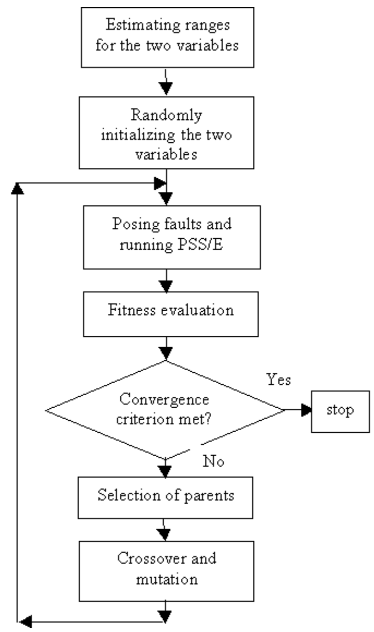

Figure 3: The flowchart for the GA based fault location estimation

In GA based optimization approach, (2) will be chosen as the fitness function. Fault location x and the fault resistance R_f are the two variables. The flow chart of the algorithm is shown in Fig. 3. The ranges for varying x and R_f can be decided as follows. R_f can be selected according to the typical possible fault resistance values. x can be simply selected as ranging from zero to the sum of the length of all the possible faulty lines, or estimated by other suitable algorithms [2, 4].

In the step “Posing faults and running PSS/E”, the short circuit studies are carried out by using the software package PSS/E according to the specified fault conditions [5]. In “Fitness evaluation”, (2) is applied to evaluate the matching degree of the simulated and recorded phasors. Steps related to initialization, selection of parents, crossover and mutation are standard GA steps, as illustrated in [3].

By iteratively posing faults, running short circuit simulations, evaluating the fitness value, updating the fault location and resistance, the GA based searching engine guides the searching process for a globally optimal solution.

IV. EVALUATION STUDIES

This section presents simulation studies utilizing the sample system depicted in Fig. 1. Results obtained by the existing approach have also been presented for comparison purposes [2]. The GA uses the following parameters: population size: 30, crossover probability: 0.85, mutation probability: 0.05, coding binary string length for fault location: 9, and coding binary string length for fault resistance: 8 [3]. Fault location ranges from 0 to the sum of the length of all the lines, as shown in Fig. 1. Fault resistance ranges from 0 to 0.4 p.u.

Results for various case studies are listed in Table 1. In comparison, results obtained by the existing algorithm assuming the fault resistance as zero are also reported [2]. In the table, x_a is the actual fault location, x_e the estimated fault location by GA, R_{fa} the actual fault resistance, R_{fe} the estimated fault resistance by GA, and x_t the fault location yielded by [2]. The second column of the table indicates the names of the buses that have DFR recordings available for fault location estimation. The sixth, ninth and twelve columns of the table indicate the actual faulty lines, and the faulty lines estimated by the GA and existing approaches respectively. In these columns, the first and second numbers represent the starting bus and the end bus of the line, respectively.

Note that the fault location listed in the table refers to the distance of the fault from the starting bus of the faulty line. For example in case 1, it is shown that the actual fault is on the line from bus 13 to 12 and is 9.1 miles away from bus 13. The estimated fault location by GA is on the line from bus 13 to 12 and is 8.8 miles away from bus 13, and the fault location obtained by [2] is 550.4 miles from bus 1. "N/A" in the last column indicates that the corresponding approach is not able to reliably identify the faulty line. It is seen that for all the cases, the GA approach is able to pinpoint the fault location quite accurately irrespective of the fault type and fault resistance, while the existing approach does not give reliable results

V. CONCLUSIONS

To improve the accuracy for fault location estimation when "sparse data" are available, the "waveform matching" concept may be utilized. A new GA based approach for implementing this concept has been introduced in this paper. Evaluation studies utilizing simulated data have shown that the proposed approach may lead to improved accuracy over existing ones for fault location when only "sparse data" are available. Software implementation issues and the test results using field data will be reported in future papers.

VI. ACKNOWLEDGEMENTS

The work presented in this paper was funded by the Texas Higher Education Coordinating Board Advanced Technology Program. The co-funding is provided by TXU Electric and Reliant Energy HL&P.

VII. REFERENCES

1. Girgis A A, Hart D G, and Peterson W L, 1992, "A new fault location technique for two- and three-terminal lines", IEEE Transactions on Power Delivery, 7, 98-107.
2. Waikar D L, Elangovan S, and Liew A C, 1994, "Fault impedance estimation algorithm for digital distance relaying", IEEE Transactions on Power Delivery, 9, 1375-1383.
3. Goldberg D E, 1989, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison Wesley, Reading, MA.
4. Kezunovic M, 1998, "Automation of fault analysis using DFR data", Proceedings of the 60th American Power Conference, Chicago, IL, pp. 91-98.
5. Power Technologies, Inc., 1997, "PSS/E 25 Program Operation and Application Manuals".

TABLE 1 - Fault location estimated by the proposed approach and the existing approach.

Case No.	Buses Having DFRs	Fault type	x_a (mile)	R_{fa} (p.u.)	Actual Faulty line	x_e (mile)	R_{fe} (p.u.)	Faulty line by GA	Iteration Times	x_t (mile)	Faulty line By [2]
1	1	a-g	9.1	0.005	13 – 12	8.8	0.003	13 – 12	25	550.4	N/A
2	1	a-g	9.1	0.4	13 – 12	10.1	0.38	13 – 12	27	-656.5	N/A
3	1	a-b	9.1	0.005	13 – 12	10.1	0.006	13 – 12	26	527.1	N/A
4	1	a-b	9.1	0.4	13 – 12	7.5	0.37	13 – 12	27	-316.3	N/A
5	1	a-b-g	9.1	0.005	13 – 12	11.4	0.003	13 – 12	28	546.3	N/A
6	1	a-b-g	9.1	0.4	13 – 12	7.5	0.43	13 – 12	25	-653.2	N/A
7	1	a-b-c	9.1	0.0	13 – 12	8.8	0.002	13 – 12	24	527.6	N/A