A Robust Estimation Method for Topology Error Identification

L. Mili G. Steeno F. Dobraca D. French

Bradley Department of Electrical Engineering
Virginia Polytechnic Institute & State University
Blacksburg, VA 24061-0111

Siemens Power Transmission & Distribution (LLC)
7225 Northland Drive
Brooklyn Park, MN 55428-1540

Abstract: A pre-processing method that identifies both multiple topology errors and bad measurements is described. The method determines the branch statuses by testing the real and reactive power flow estimates of all the branches of the network, irrespective of their assumed statuses. The power flows are the state variables of two decoupled real and reactive power models that stem from both a detailed substation representation and a super-node modeling. They are estimated by means of the iteratively reweighted least-squares algorithm that implements the Huber M-estimator. The procedure is not prone to divergence problems, which is of great value in a real-time environment. The performance of the method is demonstrated on the IEEE-118 bus system.

1. Introduction

Derived by Schweppe *et al.* [1] in the early seventies, power system state estimation is based on a super-bus network modeling in which the power and voltage measurements are expressed in terms of the system state variables, chosen to be the nodal voltage magnitudes and phase angles. In this model, the parameters of the lines and transformers along with the topology of the network are assumed to be known. The super-bus network is built by the topology processor from the statuses of the circuit breakers, which are either telemetered or determined manually by the operators. A mistake in the breaker statuses of a piece of equipment, which may be either a line, or a transformer, or a shunt capacitor, or a bus coupler, yields a topology error. This in turn translates into multiple conforming bad measurements in the state estimation model.

This property has prompted the development of several topology error identification methods conceived as post-state-estimation procedures [2-4]. Roughly speaking, the topology of the system is suspected once the measurements associated with a branch or a bus are flagged as outliers by a state-estimation-based residual test. However, these post-processing methods may fail due to the conforming nature of the bad measurements and their number, which may be very large, especially in the case of bus splitting errors. In that case, the conventional weighted least squares (WLS) estimator may diverge [5-7] and any robust estimator may not reject the outliers because of a local breakdown. Hence the need of a pre-processing method for topology error identification. This is precisely what were advocated by Singh and Glavitsch [7], who developed a rule-based technique, and by Bonanomi and Gramberg [8], who initiated a graph-theoretic search procedure.

PE-022-PWRS-0-06-1998 A paper recommended and approved by the IEEE Power System Operations Committee of the IEEE Power Engineering Society for publication in the IEEE Transactions on Power Systems. Manuscript submitted September 4, 1997; made available for printing June 12, 1998.

Another interesting approach was proposed by Irving and Sterling [9]. They suggested to carry out data validation at the substation level where all the bus couplers and the circuit breakers are represented in detail. To be able to handle zero impedance branches, they suggested to take as variables the powers flowing through the circuit breakers. Later on, Clewer *et al.* [10] extended the approach to the whole system by proposing a 4-stage iterative procedure based on a detailed substation representation together with a bus level network modeling. Unfortunately, the method involves many inter-related steps, which makes it rather complex. In addition, it is not guaranteed to converge [10].

In an attempt to decrease the complexity of the foregoing method while meeting the need of a generalized state estimation raised by Slutsker *et al.* [11,12], Monticelli [13], Alsac *et al.* [14], and Abur *et al.* [15] advocated the use of a 2-step procedure that proceeds as follows. First, a super-node-based state estimation is executed and a residual analysis is performed. In the event that the residuals associated with a branch or a bus are found to be large, then a detailed representation of the suspected substations is carried out and the state vector is expanded accordingly. Finally, the expanded state vector is estimated through either a conventional WLS estimator [13,14] or a LAV estimator [15]. However, both methods suffer from the aforementioned weaknesses inherent to any post-estimation approach.

The paper describes a pre-processing method that identifies all types of topology errors while being able to discriminate them from bad measurements. In addition, it is not prone to divergence problems, which is of great value in a real-time environment. This is achieved by applying the Huber M-estimator to two decoupled real and reactive power models that stem from both a detailed substation representation and a super-node modeling. The state variables are the collection of the real and reactive power flows located at one end of every branch of the network. The branch statuses are determined on the basis of a statistical test applied to the power flow estimates. These estimates are calculted by means of the iteratively reweighted least-squares algorithm, which exhibits good convergence rates regardless of the X/R ratios of the lines.

The paper is organized as follows. Section 2 is devoted to the network modeling. Sections 3 and 4 derive the decoupled power models, respectively. The Huber M-estimator is outlined in Section 5 and the power flow statistical test is described in Section 6. Simulation results carried out on the IEEE 118-bus system are given in Section 7.

2 - Network Modeling

The network modeling encompasses all the lines, transformers, and shunt capacitors/reactors of the system, regardless of their assumed statuses. In this model, a substation containing a bus bar that may split in two well defined segments is represented by two distinct super-nodes connected through a zero impedance branch. On the other hand, a substation containing a bus bar that may split in several different ways is modeled in detail. In the event where such substations are large in number, those that have real and reactive power injection measurements attached to them may be first modeled as super-nodes. Then, the power injections are estimated and their residuals tested. If a residual is found to be large, the corresponding substation is modeled in detail and the state re-estimated.

3 - Real Power Modeling

The intent here is to derive a model that relates the real power measurements z_{Pi} to the state variables x_{Pi}, each of which being associated with a branch of the network. In this model, every branch has a state variable associated with it according to the following rule: When a branch has a single real power flow measurement, the associated state variable is the flow of that measurement, otherwise it is either real power flow.

To fix the ideas, consider a branch that connects node k to node l. It will be represented by a pi-equivalent circuit as displayed in Fig. 1. It can be easily shown that the Ohmic losses dissipated in the branch are given by

$$P_{L_{kl}} = G_{kl}\ (V_k^2 + V_l^2 - 2\, V_k\, V_l \cos\theta_{kl}), \qquad (1)$$

where V_k and V_l are the voltage magnitudes at the nodes k and l, respectively. Here, θ_{kl} denotes the nodal voltage angle across the branch k-l, and G_{kl} is the series conductance defined as $G_{kl} = R_{kl} / (R_{kl}^2 + X_{kl}^2)$. Assume that x_{Pi} is the power flow on the k side of the branch directed toward node l as shown in Fig.1. Using the dc model, θ_{kl} may be written as the product of the state variable x_{Pi} and the series reactance X_{kl}, that is,

$$\theta_{kl} \approx X_{kl}\, x_{Pi}. \qquad (2)$$

Substituting (2) into (1) and putting $V_k = V_l = 1$ pu gives

$$P_{L_{kl}} \approx 2\, G_{kl}\ (1 - \cos(X_{kl}\, x_{Pi})). \qquad (3)$$

As an alternative, V_k and V_l may be set equal to the corresponding metered values. Surprisingly, formula (3) is found to be accurate enough in estimating the branch Ohmic losses regardless of the line X/R ratios (see the remark given in Section 6.)

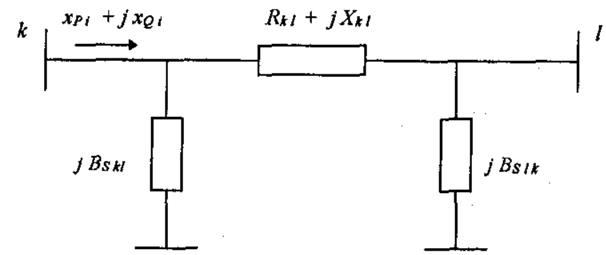

Fig.1 . Pi-equivalent circuit of the branch k-l.

By making use of (3), the power flows P_{kl} and P_{lk} and the power injections P_k and P_l may be written as

$$P_{kl} = x_{Pi}, \qquad (4)$$

$$P_{lk} = - x_{Pi} + P_{L_{kl}} = - x_{Pi} + 2\, G_{kl}\, (1 - \cos(X_{kl}\, x_{Pi})), \qquad (5)$$

$$P_k = \sum_{l \in N(k)} P_{kl}\ , \qquad P_l = \sum_{k \in N(l)} P_{lk} \qquad (6)$$

where $N(k)$ and $N(l)$ denotes the set of nodes adjacent to node k and l, respectively.

Suppose now that the network has n_b state variables contained in the vector $\underline{x}_P$ and is provided with m_P real power measurements contained in the vector $\underline{z}_P$. This leads to a nonlinear regression model that is written as

$$\underline{z}_P = \underline{h}_P(\underline{x}_P) + \underline{e}_P, \qquad (7)$$

where $\underline{e}_P$ is an m_P-dimensional error vector assumed to be a random vector with mean zero and a known covariance matrix, $\underline{R}_P = diag(\sigma_{P1}^2, \ldots, \sigma_{Pmp}^2)$. In (7), the state vector $\underline{x}_P$ may be estimated by means of the iteratively reweighted least-squares (IRLS) algorithm that solves the Huber estimator. The approach will be described in Section 5. It makes use of the Jacobian matrix $\underline{H}_P(\underline{x}_P) = \partial \underline{h}_P(\underline{x}_P) / \partial \underline{x}_P$, whose entries are the partial derivatives of P_{kl}, P_{lk}, and P_k with respect to x_{Pi}. They are given by

$$\frac{\partial P_{kl}}{\partial x_{Pi}} = 1, \qquad (8)$$

$$\frac{\partial P_{lk}}{\partial x_{Pi}} = -1 + 2\, X_{kl}\, G_{lk} \sin(X_{kl}\, x_{Pi}), \qquad (9)$$

$$\frac{\partial P_k}{\partial x_{Pi}} = \sum_{l \in N(k)} \frac{\partial P_{kl}}{\partial x_{Pi}}. \qquad (10)$$

4 - Reactive Power Modeling

The model for the reactive powers is similar to the real power model in that every branch has a state variable associated with it. This state variable is defined as one of the two reactive power flows of the branch according to the following rule: When a branch is provided with a single reactive power flow measurement, the state variable is the associated flow, otherwise it is either reactive flow.

Consider now the branch *k-l* depicted in Fig. 1. The state variable x_{Qi} for the reactive model is assumed to be the reactive power flowing from node k to node l. The reactive powers on both ends of the branch can then be expressed as

$$Q_{kl} = x_{Qi}, \tag{11}$$

$$Q_{lk} = -x_{Qi} + Q_{L_{kl}}, \tag{12}$$

and

$$Q_k = \sum_{l \in N(k)} Q_{kl}, \quad Q_l = \sum_{k \in N(l)} Q_{lk}, \tag{13}$$

where $N(k)$ and $N(l)$ are defined as in (6), and $Q_{L_{kl}}$ is the reactive power consumed by the branch. It is given by

$$Q_{L_{kl}} = -B_{S_{kl}}(V_k^2 + V_l^2) - B_{kl}(V_k^2 + V_l^2 - 2V_k V_l \cos\theta_{kl}). \tag{14}$$

Similarly to the Ohmic losses of the branch, we substitute in (14) the expression of θ_{kl} given by (2) and we put $V_k = V_l = 1$ pu, yielding

$$Q_{L_{kl}} \approx -2B_{S_{kl}} - 2B_{kl}(1 - \cos(X_{kl} x_{Pi})). \tag{15}$$

Here too, V_k and V_l may be put equal to the metered values instead, if they are available. We notice that, once an estimate $\hat{x}_{Pi}$ for the real power flow x_{Pi} is calculated, an estimate $\hat{Q}_{L_{kl}}$ for $Q_{L_{kl}}$ is readily obtained through

$$\hat{Q}_{L_{kl}} \approx -2B_{S_{kl}} - 2B_{kl}(1 - \cos(X_{kl} \hat{x}_{Pi})). \tag{16}$$

Replacing $Q_{L_{kl}}$ by $\hat{Q}_{L_{kl}}$ in (12) yields

$$Q_{lk} \approx -x_{Qi} + \hat{Q}_{L_{kl}}. \tag{17}$$

Interestingly, $\hat{Q}_{L_{kl}}$ is found to be accurate enough for topology error identification (see the remark given in Section 6). Since it is independent of the state variables x_{Qi}, it can be calculated just after the estimation of the real power flows, x_{Pi}.

We are now ready to derive the reactive power model. To do that, let us assume that the network is provided with m_Q reactive power measurements and n_b state variables. Let $\underline{z}_Q$ and $\underline{x}_Q$ denote the measurement vector and the state vector, respectively. The reactive power model in matrix form is then given by

$$\underline{z}_Q = \underline{h}_Q(\underline{x}_Q) + \underline{e}_Q, \tag{18}$$

where $\underline{e}_Q$ is an m_Q-dimensional error vector assumed to be a random vector with mean zero and a known covariance matrix, $\underline{R}_Q = diag(\sigma_{Q1}^2, \ldots, \sigma_{Qm_Q}^2)$. Similarly to $\underline{x}_P$ and as described in Section 5, the state vector $\underline{x}_Q$ is estimated by means of the IRLS algorithm that solves the Huber estimator. It makes use of the Jacobian matrix $\underline{H}_Q(\underline{x}_Q) = \partial \underline{h}_Q(\underline{x}_Q) / \partial \underline{x}_Q$ whose entries are the partial derivatives of Q_{kl}, Q_{lk}, Q_k, and Q_l with respect to x_{Qi}. They are expressed as

$$\frac{\partial Q_{kl}}{\partial x_{Qi}} = -\frac{\partial Q_{lk}}{\partial x_{Qi}} = 1, \tag{19}$$

and

$$\frac{\partial Q_k}{\partial x_{Qi}} = \sum_{l \in N(k)} \frac{\partial Q_{kl}}{\partial x_{Qi}}. \tag{20}$$

5 - The Huber Estimator

By dropping the subscript P and Q, the real and the reactive power models given by (7) and (18), respectively, may be written as

$$\underline{z} = \underline{h}(\underline{x}) + \underline{e}. \tag{21}$$

The state vector $\underline{x}$ is estimated by means of the Huber estimator, which minimizes an objective function given by

$$J(\underline{x}) = \sum_{i=1}^{m} \rho(r_{wi}), \tag{22}$$

where

$$\rho(r_{wi}) = \begin{cases} \frac{1}{2} r_{wi}^2 & |r_{wi}| \le \lambda \\ \lambda |r_{wi}| - \frac{\lambda^2}{2} & elsewhere \end{cases} \tag{23}$$

Here, $r_i = z_i - h_i(\underline{x})$ is the ith residual, $r_{wi} = r_i / \sigma_i$ is the ith weighted residual, and λ is a cutoff value fixed at 2.7. The estimate $\hat{\underline{x}}$ is a solution to $\partial J(\underline{x}) / \partial \underline{x} = \underline{0}$, namely to

$$\sum_{i=1}^{m} \underline{\ell}_i \; \psi(r_{wi}) = \underline{0}. \tag{24}$$

In (24), $\underline{\ell}_i^T$ is the ith row of the weighted Jacobian matrix, $\sqrt{\underline{R}^{-1}}\ \underline{H}$, and $\psi(r_{wi})$ is the derivative of $\rho(r_{wi})$ with respect to r_{wi}. It is given by

$$\psi(r_{wi}) = \begin{cases} r_{wi} & |r_{wi}| \le \lambda \\ \lambda & elsewhere \end{cases} \quad . \qquad (25)$$

To solve (24) for $\underline{x}$, let us divide and multiply $\psi(r_{wi})$ by r_{wi}. This gives

$$\sum_{i=1}^{m} \frac{1}{\sigma_i} \underline{\ell}_i \ r_i \ q(r_{wi}) = \underline{0}, \qquad (26)$$

where $q(r_{wi}) = \psi(r_{wi})/r_{wi}$ is called the weight function. Defining $\underline{Q} = diag(q(r_{wi}))$ and putting (26) in a matrix form yields

$$\underline{H}^T \ \underline{R}^{-1} \ \underline{Q} \ (\underline{z} - \underline{h}(\underline{x})) = \underline{0}. \qquad (27)$$

Substituting in (27) $\underline{h}(\underline{x})$ by a first order Taylor series expansion around $\underline{x}^{(k)}$, that is, by

$$\underline{h}(\underline{x}) = \underline{h}(\underline{x}^{(k)}) + \underline{H}(\underline{x}^{(k)})(\underline{x} - \underline{x}^{(k)}), \qquad (28)$$

and rearranging the terms of the equation so obtained, we get the so-called iteratively re-weighted least squares (IRLS) algorithm. It is written as

$$\Delta\underline{x}^{(k)} = (\underline{H}^{(k)T} \ \underline{R}^{-1} \ \underline{Q}^{(k)} \ \underline{H}^{(k)})^{-1} \ \underline{H}^{(k)} \ \underline{R}^{-1} \ \underline{Q} \ \underline{r}^{(k)}, \qquad (29)$$

where $\Delta\underline{x}^{(k)} = \underline{x}^{(k+1)} - \underline{x}^{(k)}$. Unlike a Newton-type algorithm that involves divisions by the second derivative of the $\rho(.)$ function, which may vanish, the IRLS algorithm proved to be numerically stable while exhibiting good convergence properties.

Once the state vector is estimated, the extreme outliers are identified and deleted from the measurement set. Then, one iteration of the IRLS algorithm is executed starting from the previous solution. This measurement deletion is performed in order to cancel out the influence of the extreme outliers on the estimates. They are defined as measurements whose weighted residual have an amplitude larger than a given threshold, chosen between 6 and 10.

When solving (29), most of the computing time is spent in factorizing the matrix $\underline{G} = \underline{H}^T \underline{R}^{-1} \underline{Q} \ \underline{H}$. A good scheme would be to factorize $\underline{G}$ at the first iterations and then to carry out partial matrix refactorization based on the sparse vector method as advocated by Chan and Brandwajn [16]. This scheme takes advantage of the fact that only few diagonal elements of the $\underline{Q}$ matrix usually change after two or three iterations.

6 - Testing the Power Flow Estimates

Once the power flows are estimated, their amplitudes are tested and the statuses of the associated branches determined. Here, a robust statistical test is applied on each branch to the pair of P and Q flow estimates that are largest in amplitude. This increases the power of the statistical test, which in turn decreases the number of cases where a branch is wrongly tagged as being disconnected.

The statistical test is carried out as follows. Let $\hat{\underline{z}}$ denote the vector of all the flow estimates, be they associated with a measurement or not. The amplitudes of the standardized flow estimates,

$$\hat{z}_i \ / \ \hat{s}_{zi}, \qquad (30)$$

is compared to a given cutoff value, say 3. If an amplitude is found to be smaller than 3, then the associated branch is labeled off, otherwise it is labeled on. The scale estimates $\hat{s}_{zi}$ are the square-root of the diagonal entries of the covariance matrix of $\hat{\underline{z}}$. Let us derive its expression. To do that, we first write $\hat{\underline{z}}$ as a function of the state vector $\hat{\underline{x}}$, that is,

$$\hat{\underline{z}} = \underline{\tilde{h}}(\hat{\underline{x}}). \qquad (31)$$

Then, we perform a first order Taylor series expansion of $\underline{\tilde{h}}(\underline{x})$ about $\hat{\underline{x}}$ to get

$$\underline{\tilde{h}}(\underline{x}) = \underline{\tilde{h}}(\hat{\underline{x}}) + \underline{\tilde{H}}(\hat{\underline{x}})(\underline{x} - \hat{\underline{x}}). \qquad (32)$$

Substituting (32) into (31) yields

$$\hat{\underline{z}} = \underline{\tilde{h}}(\underline{x}) - \underline{\tilde{H}}(\hat{\underline{x}})(\underline{x} - \hat{\underline{x}}). \qquad (33)$$

Hence, the covariance matrix of $\hat{\underline{z}}$ is written as

$$Cov(\hat{\underline{z}}) = \underline{\tilde{H}}(\hat{\underline{x}}) \, Cov(\hat{\underline{x}}) \, \underline{\tilde{H}}^T(\hat{\underline{x}}). \qquad (34)$$

In (34), the covariance matrix of $\hat{\underline{x}}$, $Cov(\hat{\underline{x}})$, has been derived by Huber [17] in the asymptotic case, namely, when the number of measurements tend to infinity. It is given by

$$Cov(\hat{\underline{x}}) = \hat{s}_x \ (\underline{H}^T \ \underline{R}^{-1} \ \underline{H})^{-1}, \qquad (35)$$

where $\hat{s}_x$ is an appropriate scale estimate. Under Gaussian measurement errors, Monte Carlo simulations carried out on various test systems showed that $\hat{s}_x$ needs to be put equal to 1.2.

Note that only the diagonal elements of the flow estimate covariance matrix given by (34) are needed. A very efficient way to calculate them is to use the sparse vector method [18].

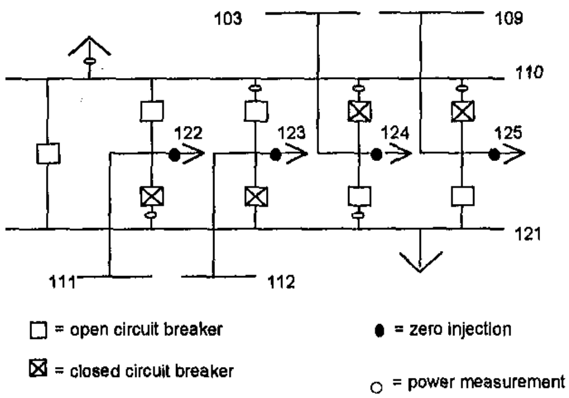

Fig 2. Substation modeling associated with Bus 2.

Remark: It is true that for a heavily loaded line with a low X/R ratio, the accuracy of the P and Q loss estimates given by (3) and (16) may deteriorate. However, this has no consequence on the results of the statistical test as applied to the line power flow estimates. These estimates will be significantly different from zero and hence, the line will be labeled on as it should be.

7 - Simulation Results

The topology estimator has been applied to a 120-bus system derived from the IEEE 118-bus system by assuming that two pairs of buses, namely Buses 11 and 119 and Buses 24 and 120, are each connected by a bus coupler that may split them in only one way. Consequently, they are not merged in two super-nodes. As for the other buses, they are each represented by a super-node because they have either no bus coupler or have one that may split them in many different ways. The measurement configuration consists of 240 pairs of P and Q measurements along with 19 pairs of zero injections. They will be used to estimate 181 pairs of power flows, the state variables of the P and Q models.

Several cases of topology errors and bad measurements have been simulated. Only one case will be described next. In all simulated cases, the IRLS algorithm converges in 4 to 7 iterations for the P model, and in 2 to 6 iterations for the Q model. Here, the iterations are stopped whenever the absolute incremental changes in x_i are less than 0.05 pu for all state variables if the number of iterations k does not exceed 5, and for only those metered state variables with $|r_i/\sigma_i| \leq 2.0$ if $k > 5$.

We introduce 9 topology errors and 8 bad measurements in the system. Specifically, we assume that Lines 15-19, 49-54, and 69-70 are disconnected while they are connected and that Lines 27-115, 28-29, and 62-67 are connected while they are disconnected. We also assume that the bus coupler between Buses 11 and 119 is closed while it is open and that the bus coupler between Buses 24 and 120 is open while it is closed. Finally, we suppose that the bus couplers located at the substation of the super-node Bus 110 are switched on or off so that we have not one, but two separate buses. Now, since Bus 110 is represented as a super-node, we expect that the real and reactive power injection measurements at that bus be rejected as bad data by the Huber estimator. It does just that by providing these injection measurements with weighted residuals of -7.61 and 5.56, respectively. In the second step of the procedure, the substation associated with Bus 110 is represented in detail as shown in Fig. 2. This leads to a system with 125 buses and 190 branches. The Huber estimator is then executed and the weighted residuals calculated. They are displayed in Table 1 for the 8 bad measurements. We observe that all of them have been properly rejected with large absolute weighted residuals. The extreme bad data with $|r_i/\sigma_i| \geq 10$ are then deleted, one iteration of the IRLS algorithm is carried out, and a statistical test is applied to the flow estimates. It is found that the test identifies correctly the statuses of all the branches. For instance, it is seen in Table 2 that Branches 110-121, 110-122, 110-123, 121-124, and 121-125 are tested off, revealing that Bus 110 splits in two separate buses, as it should be.

Table 1. Weighted residuals of the bad measurements.

Measurement	*Exact value*	*Est. value*	*Weight. Resid.*
PFL 30-8	-0.72	-0.66	30.52
PFL 49-69	-0.48	-0.59	26.82
PFL 63-59	1.55	1.49	-42.13
PIN 115	-0.22	-0.07	-13.80
QFL 30-8	-0.73	-0.66	20.88
QFL 49-69	0.11	-0.07	-17.54
QFL 63-59	-0.34	-0.33	30.23
QIN 115	-0.07	0.08	-11.09

Table 2. Standardized flows and branch statuses.

Branch	$\hat{P}_{kl}$	$\hat{P}_{kl}/\hat{s}_{zi}$	$\hat{Q}_{kl}$	$\hat{Q}_{kl}/\hat{s}_{zi}$	*Status*
15-19	0.14	6.60	-0.17	-4.52	ON
49-54	0.71	23.07	-0.27	-5.37	ON
70-69	1.02	77.72	0.16	6.75	ON
27-115	0.01	0.59	-0.02	-1.06	OFF
28-29	0.01	1.17	-0.02	-1.00	OFF
62-67	0.01	0.57	-0.02	-1.30	OFF
11-119	0.28	22.75	-0.34	-14.32	ON
24-120	0.02	1.31	-0.01	-0.52	OFF
110-121	0.11	1.23	0.05	0.55	OFF
110-122	0.01	0.15	0.00	0.08	OFF
121-122	0.36	28.16	0.00	0.06	ON
110-123	0.04	0.66	-0.02	-0.29	OFF
121-123	0.68	54.23	-0.30	-11.90	ON
110-124	0.58	56.69	0.03	1.21	ON
121-124	0.01	0.86	-0.01	-0.41	OFF
110-125	0.13	10.63	0.10	4.11	ON
121-125	0.01	0.21	0.01	0.16	OFF

8 - Conclusions

The developed method is a pre-processing procedure that robustly estimates the topology of the system in presence of both multiple topology errors and bad measurements. System topology is determined by testing the real and reactive power flows of all the branches of the network after being estimated by means of the IRLS algorithm that implements the Huber M-estimator. The algorithm can be made compatible with real-time environment through advanced sparsity techniques and numerically stable methods such as sparse vector method, partial matrix factorization, and Givens rotations.

9 - Acknowledgements

Financial support provided by Siemens Power Transmission and Distribution and by NSF under grant ECS-9257204 is gratefully acknowledged.

10 - References

[1] F.C. Schweppe, J. Wildes, and D.B. Rom, "Power System Static State Estimation", Part I, II, III, *IEEE Trans. on Power Apparatus and Systems*, Vol. PAS-89, No. 1, Jan. 1970, pp. 120-135.

[2] K. A. Clements and P. W. Davis, "Detection and Identification of Topology Errors in Electric Power Systems," *IEEE Transactions on Power Systems*, Vol. 3, No. 4, Nov. 1988, pp. 1748-1753.

[3] F. F. Wu and W.-H. E. Liu, "Detection of Topology Errors by State Estimation," *IEEE Transactions on Power Systems*, Vol. 4, No. 1, Feb. 1989, pp. 176-183.

[4] A. S. Costa and J. A. Leao, "Identification of Topology Errors in Power System State Estimation," *IEEE Transactions on Power Systems*, Vol. 8, No. 4, Nov. 1993, pp. 1531-1538.

[5] M. Assadian, R. J. Goddard, H. W. Hong, and D. French, "Field Operational Experiences with on Line State Estimation," *IEEE Transactions on Power Systems*, Vol. 9, No. 1, Feb. 1994, pp. 50-58.

[6] R. A.M. van Amerongen, "On Convergence Analysis and Convergence Enhancement of Power System Least-Squares State Estimators," *IEEE Transactions on Power Systems*, Vol. 10, No. 4, Nov. 1995, pp. 2038-2044.

[7] N. Singh and H. Glavitsch, "Detection and Identification of Topological Errors in Online Power System Analysis," *IEEE Transactions on Power Systems*, Vol. 6, No. 1, Feb. 1991, pp. 324-331.

[8] P. Bonanomi and G. Gramberg, "Power System data Validation and State Calculation by Network Search Techniques," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-102, No. 1, Jan. 1983, pp. 238-249.

[9] M. R. Irving, M. J. H. Sterling, "Substation Data Validation," IEE Proceedings, Part C, Vol. 129, No. 3, May 1982, pp. 119-122.

[10] B. C. Clewer, M. R. Irving, and M. J. H. Sterling, "Topologically Independent State Estimation," *Proceedings of the IFAC Symposium on Power Systems, Modeling and Control Applications*, Brussels, 5-8 Sept. 1988, pp. 17.1.1.- 17.1.13.

[11] I. Slutsker and S. Mokhtari, "Comprehensive Estimation in Power Systems: State, Topology, and Parameter Estimation," Proceedings of the American Power Conference, Vol. 57-I, Chicago, Illinois, April 1995, pp. 149-155.

[12] I. Slutsker, M. Mirheydar, M. Cahagan, H.-T. Neisius, G. Schellstede, "Scan Rate Estimation in Energy Management Systems," Proceedings of the American Power Conference, Chicago, Illinois, April 1996, pp. 924-929.

[13] A. Monticelli, "Modeling Circuit Breakers in Weighted Least Squares State estimation," *IEEE Transactions on Power Systems*, Vol. 8, No. 3, Aug. 1993, pp. 1143-1149.

[14] O. Alsac, N. Vempati, B. Stott, and A. Monticelli, "Generalized State Estimation," Proceedings of the PICA Conference, May 11-16, 1997, Columbus, OH, pp. 90-96.

[15] A. Abur, H. Kim, and M. K. Celik, "Identifying the Unknown Circuit Breaker Statuses in Power Networks", *IEEE Trans. on Power Systems*, Vol. 10, No. 4, Nov. 1995, pp. 2029-2037.

[16] S. M. Chan and V. Brandwajn, "Partial Matrix Factorization," *IEEE Transactions on Power Systems*, Vol. PWRS-1, No. 1, Feb. 1986, pp. 193-200.

[17] P.J. Huber, *Robust Statistics*. John Wiley, 1981.

[18] W.F. Tinney, V. Brandwajn, and S.M. Chan, "Sparse Vector Method," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-104, No. 2, February 1985, pp. 295-301.

Lamine Mili received a Electrical Engineering diploma from EPFL, Switzerland, in 1976, and a Ph.D. from the University of Liège, Belgium, in 1987. He is presently an Associate professor of Electrical Engineering at Virginia Tech. His research interests are in power system analysis and control.

Gregory Steeno received his B.S. in Statistics in 1994 and his M.S. in Statistics in 1995, both from Virginia Tech. He is currently working toward his Ph.D. degree at Virginia Tech. His research interests include robust estimation theory, nonparametric statistics, and power system state estimation.

Fadil Dobraca received his B.S.E.E. from the University of Sarajevo, Bosnia-Herzegovina, his M.S. and Ph.D. in Electrical Engineering from the University of Illinois at Urbana-Champaign. He is with Siemens Power Transmission & Distribution, involved in network applications.

Dan French received his BSEE from the Illinois Institute of Technology, in 1979, and his MSEE from the University of Illinois-Urbana, in 1981. He is with Siemens Power Transmission & Distribution, involved in network applications.

Discussion

Hongrae Kim (SoonChunHyang University, KOREA): The authors are to be congratulated for their paper on topology error identification in electric power systems. This paper presents an estimation/topology identification method by implementing IRLS algorithm and Huber estimator. I have a number of questions on which the authors views would be appreciated.

1. In section 2, the authors say that a substation containing a bus bar that may split in several different ways is modeled in detail. If the system is modeled in detail, the dimension of the system becomes large and it takes much time to compute the estimates. Also, for the detailed model the measurements may not be locally redundant. How could the authors solve these problems?
2. The proposed method is a pre-processing procedure for identifying topology errors. In order to apply this method to real systems, it should be fast enough to identify topology errors and calculate the states. Please provide the timings.
3. The authors set a threshold value(between 6 and 10) in section 5, a cutoff value(3) in section 6, a convergence tolerance(0.05 pu) in section 7, and some others. What are the bases for these values?
4. The authors show one simulated case. In the case, there are several branch topology errors and one bus split error being split in two separate buses. Bus split can be happened in many different ways according to the statuses of the bus couplers and bus sections. Could the proposed method deal with all the possible bus split cases?

L. Mili, G. Steeno, F. Dobraca, and D. French: We thank Dr. Kim for his interest in our paper. We will answer to the questions in the order that they have been raised:

1. As indicated in the paper, we do recommend modeling a substation in detail whenever the bus bar may split in several different ways. However, this detailed representation requires the availability of enough measurements at the substation level so that the system remains observable. In addition, it will increase significantly the size of the system, and thereby the number of state variables to be estimated. We have been working on an observability algorithm, as well as an algorithm for the addition of pseudo-measurements to increase the measurement redundancy. For example, if the statuses of the circuit breakers of a line are telesignaled or assumed to be open, then we can replace that information by zero P and Q pseudo-measurements associated with that line. As additional information, we may apply rule-based voting techniques to the statuses of the circuit breakers and switches of a line and convert them to pseudo-measurements as appropriate. If the system can not be modeled in detail due to observability constraints, then we suggest supernode modeling where necessary as shown in the example of Section 7. The rejection of the power injection measurements by the estimator indicates a gross error either in the measurements or in the supernode modeling.

2. For the same number of state variables, the topology estimation algorithm has the same order of complexity as that of the SHGM state estimator proposed in [A] and based on a conventional decoupled model involving the nodal voltage magnitudes and phase angles. Since the SHGM estimator exhibits a linear growth with the number of state variables when sparsity techniques are used, we may do the following. We compare the computing time of the topology estimator with that of the SHGM estimator through the ratio of the number of state variables involved in the topology estimator model with respect to the classical state estimator. These ratios are provided in the third and fifth columns of Table A for various percentages of the substations modeled in detail, which are displayed in the first column. As a rule, we assume that each substation on average has ten lines, which implies 20 state variables per substation in addition to the 20 and 178 branches of the 14-bus system and the 118-bus system, respectively. The total number of state variables for each case is given in the second and fourth columns of Table A. We observe that the ratios do not exceed 12 when all the substations are modeled in detail. These ratios are found to be in the same range for both systems. This is probably due to the applied rule of ten lines per substation.

% of Modeled Substations	IEEE 14-Bus		IEEE 118-Bus	
	State Variables	Ratio	State Variables	Ratio
0	40	1.48	356	1.51
10	60	2.22	596	2.54
20	100	3.73	836	3.56
30	120	4.44	1056	4.49
40	160	5.92	1296	5.51
50	180	6.67	1536	6.54
60	200	7.41	1776	7.56
70	240	8.89	2016	8.58
80	260	9.63	2236	9.51
90	300	11.11	2476	10.54
100	320	11.85	2716	11.56

3. In Section 5, we set a threshold value of $\lambda = 2.7$ for the Huber estimator to make the asymptotic variances of the flow estimates nearly as small as those of the weighted

least squares (WLS) estimates under Gaussianity of the measurement errors; the ratios being approximately equal to 1.001. However, any value of λ between 1.0 and 3.0 might have been chosen as well because it is simply used by the Huber estimator to smoothly downweight deviant measurements from Gaussianity. It is not a dichotomous decision, to keep or delete, as it is the case with the WLS bad data processing methods.

Also described in Section 5 is a measurement deletion scheme that cancels out possible influence of extreme deviant measurements on the power flow estimates, which show up with high weighted residuals. These measurements may be bad data. By deleting them and carrying out one iteration of the IRLS algorithm from the previous solution, we obtain less biased power flow estimates. We suggest a larger cutoff value, say between 6 and 10, for two reasons: First, we do not want to delete too many measurements and thereby, possibly making the network unobservable. Second, we increase the chance to delete truly bad data.

In Section 6, we set a threshold value of 3.0 for testing whether the sending end power flow estimates of a branch are significantly large and hence, to decide whether that branch is energized. This is based on the fact that when a branch is not energized, each of these estimates is asymptotically Gaussian with zero mean and unit variance. Note that even though a test may falsely declare a branch with little power as off, it will have no significant effect on the classical state estimation solutions.

In Section 7, we chose a stopping criterion of 0.05 pu for all changes in power flow estimates between one iteration and the next. Based on simulation studies, we felt this criterion is good enough for estimates associated with measurements having small weighted residuals, say smaller than 2.0. To help the convergence of the algorithm, this stopping criterion is no longer applied after 5 iterations to the estimates associated with discordant measurements.

4. The method can handle very easily all possible bus split cases as shown in the example in Section 7. The only requirement is that the station is modeled in detail and there are enough measurements to observe the system.

References

[A] L. Mili, M. G. Cheniae, N. S. Vichare, P. J. Rousseeuw, "Robust State Estimation Based on Projection Statistics," *IEEE Transactions on Power Systems*, Vol. 11, No. 2, May 1996, pp. 1118-1127.

Manuscript received February 25, 1999.