Power Engineering Letters

Robust State Estimation Using Mixed Integer Programming

M. R. Irving

Abstract—**This letter describes a robust state estimator based on the solution of a mixed integer program. A tolerance range is associated with each measurement and an estimate is chosen to maximize the number of estimated measurements that remain within tolerance (or equivalently minimize the number of measurements out of tolerance). Some small-scale examples are given which suggest that this approach is robust in the presence of gross errors, is not susceptible to leverage points, and can solve some pathological cases that have previously caused problems for robust estimation algorithms.**

***Index Terms*—Mathematical programming, robustness, state estimation.**

I. Introduction

ROBUST estimation algorithms are required in power systems to allow accurate estimation of the operating state when the measurement set is corrupted by one or more gross errors. Various robust estimators have been proposed, including weighted least squares with hypothesis testing, iteratively re-weighted least squares, weighted least absolute values, least median of squares, least trimmed squares, and various others [1], [4], [5]. A different approach has been proposed by the present author and has been implemented in reference [2]. This approach associates a tolerance range (or uncertainty range) with each measurement. So for example, if we have a power flow measurement of 103 MW, based on our knowledge of the accuracy of the metering system, we might define a tolerance range of $+1.5$ MW to -2.5 MW. The implication being that if the measurement is "good," then estimated flow should lie within the stated range (100.5 to 104.5). An estimated flow outside this range would imply that the measurement is "suspect." The proposed estimation principle is then very simple: we search for a state estimate that minimizes the number of measurements that are regarded as suspect (or equivalently, maximizes the number of estimated measurements within their tolerance).

In [2], this formulation was termed "maximum constraint satisfaction" and the problem was solved using a genetic algorithm. Here, the solution is obtained via mathematical programming and some further insight into the behavior of the method is presented.

Manuscript received February 25, 2008. Paper no. PESL-00017-2008.

The author is with the Brunel Institute of Power Systems, Brunel University, Uxbridge UB8 3PH, U.K.

Digital Object Identifier 10.1109/TPWRS.2008.926721

II. Problem Formulation

Each measurement can be represented by a pair of inequality constraints based on the upper and lower limits of tolerance for that measurement

$$h_i(\underline{x}) < z_i + t_i^+ \tag{1}$$

$$h_i(\underline{x}) > z_i - t_i^- \tag{2}$$

where

z_i	measurement i;
h_i	measurement equation i;
t_i^+	upper tolerance for measurement i;
t_i^-	lower tolerance for measurement i;
$\underline{x}$	state vector.

If all the measurement errors were within the allowable tolerance range, it would be possible to find a solution that satisfies the above inequalities for all measurements. However, if some measurements have unexpectedly large errors, it may be necessary to violate the corresponding measurement inequalities in order to find an estimate $\underline{x}$. Inequalities (1) and (2) can therefore be generalized by including $0/1$ binary variables to effectively "switch off" the inequality when necessary

$$h_i(\underline{x}) < z_i + t_i^+ + Mb_i \tag{3}$$

$$h_i(\underline{x}) > z_i - t_i^- - Mb_i \tag{4}$$

where

b_i	$0/1$ binary variable for measurement i;
M	arbitrarily large positive scalar value.

The binary variable b_i is expected to be 0 for "good" measurements. When necessary, a measurement can be ignored, or "switched off," by choosing $b_i = 1$. The scalar value M is chosen to be large enough to eliminate any influence on the solution from switched-off measurements. It is not necessary to be able to switch off the upper limit and the lower limit independently, and therefore, only a single b_i variable is defined for each measurement.

The criteria for estimation is to select a state estimate $\underline{x}$ which necessitates as few measurements as possible being ignored, i.e.

$$\text{Min.} \sum_{i=1}^{n} b_i \tag{5}$$

where n is the number of measurements.

TABLE I
ESTIMATES FOR MILI *ET AL.* EXAMPLE [4] AT VARIOUS TOLERANCES

t_i^+ (= t_i^-)	x_1	x_2	b_i
0.1	0.0197	0.4898	0 0 0 0 0 1 1
0.2	-0.1136	0.5473	0 0 0 0 0 1 1
0.3	0.0174	0.4883	0 0 0 0 0 1 1
0.4	-0.0724	0.4961	0 0 0 0 0 1 1
0.425	-0.1979	0.5322	0 0 0 0 0 1 1
0.45	0.8742	0.0624	0 0 0 1 1 0 0
0.5	0.9168	0.0492	0 0 0 1 1 0 0
0.75	1.1253	0.0468	0 0 0 1 1 0 0

TABLE II
ESTIMATES FOR RYAN EXAMPLE [5] AT VARIOUS TOLERANCES

t_i^+ (= t_i^-)	x_1	x_2	b_i
0.5	-15.546	1.2947	1 1 0 0 0 0 0 1 1
1.0	4.2333	-0.0863	0 0 0 0 1 1 1 1 0
2.0	4.5298	-0.0316	0 0 0 0 0 1 1 0 0
3.0	4.1535	-0.0390	0 0 0 0 0 1 1 0 0
4.0	4.4321	-0.0092	0 0 0 0 0 1 1 0 0
4.5	0.6364	0.4178	0 0 0 0 0 0 0 1 1
5.0	1.9286	0.3355	0 0 0 0 0 0 0 0 1

This formulation is a nonlinear mixed integer program in the binary variables b_i and the real-valued components of $\underline{x}$.

Numerical experience to date is based on linear estimation examples, and in such cases, the formulation is a mixed integer linear program. The problem can be readily solved using standard mathematical programming software, such as that available via the NEOS web-service [3].

III. RESULTS AND ANALYSIS

The robustness of the proposed method has been tested on a number of small-scale linear problems, such as the example proposed by Mili *et al.* [4], which is designed to illustrate the problems caused for many estimators by "leverage points." These are measurements that have an undue influence on the solution, due to their relatively greater distance from other measurements in the factor space [1]. Bad leverage points are particularly difficult as they are hard to detect and reject. Table I shows the results obtained with the proposed estimator on this example with two state variables and seven measurements (of which five are good measurements and two are bad leverage points). Mili *et al.* [4] show that various estimators fail on this example, giving estimates that fit the bad measurements.

Table I shows that for tolerance ranges up to ± 0.425, the proposed method rejects the last two measurements $(b_6 = b_7 = 1)$, which are the bad leverage points. Very high tolerance ranges (± 0.45 and above) allow the bad leverage points to influence the estimation. Note that the state variables, here x_1 and x_2, are referred to as parameters a and b in [4].

A further illustrative example has been introduced by Ryan [5], which includes some good measurements that happen to be co-linear with the bad measurements. This can cause least-median-of-squares estimators to "accept" the bad data points, producing anomalous estimation results. This example has two state variables and nine measurements. There are two bad measurements and three of the good measurements happen to be co-linear with these.

Table II shows that for a wide range of tolerances, from ± 1.0 to ± 4.0, the proposed method is able to reject the bad data measurements $(b_6 = b_7 = 1)$.

A small linear power network example (dc model) has been presented by Abur and Exposito [1, p. 133]. This is a four-bus network with three state variables and nine measurements. Two of the measurements are leverage points. Experience with various combinations of measurement errors has shown that the proposed algorithm can identify bad data simultaneously on both leverage points, provided the gross errors are sufficiently large (e.g., simultaneous sign errors on the measurement of flow 1-2 and injection at bus 1). Smaller errors can be more difficult to identify, however.

A limitation of the proposed method is that an estimate is found which fits the stated tolerance ranges, but there is little or no noise filtering effect among the good measurements. This limitation could be overcome by applying a two-stage process of applying the proposed estimator to identify and eliminate bad data, followed by least squares estimation on the good measurements only.

Further work is required to investigate the application to full scale ac estimation problems, but it may be noted that mixed integer programs with thousands of variables can now be solved routinely in less than one minute [3].

IV. CONCLUSION

A new approach to robust estimation has been investigated and is shown to work well for some small-scale test problems, which have proven difficult for some previous methods. Further work is needed to test the feasibility of the approach on large-scale problems.

REFERENCES

[1] A. Abur and A. G. Exposito, *Power System State Estimation Theory and Implementation*. New York: Marcel Dekker, 2004, ISBN 0-8247-5570-7.

[2] A. K. Al-Othman and M. R. Irving, "Robust state estimator based on maximum constraints satisfaction of uncertain measurements," *Measurement*, vol. 40, pp. 347–359, 2007.

[3] NEOS, NEOS Sever for Optimisation, Jan. 2008. [Online]. Available: http://neos.mcs.anl.gov/neos.

[4] L. Mili, V. Phaniraj, and P. J. Rousseeuw, "Least median of squares estimation in power systems," *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 511–523, May 1991.

[5] T. P. Ryan, *Modern Regression Methods*. New York: Wiley, 1997.