LOGIC BASED ERROR DETECTION ALGORITHM IN A SUBSTATION AUTOMATION SYSTEM

Seung Ho Hyun, Gab Ju Hwang, Gwang Won Kim, Bo gun Jin* and Seung Jae Lee*

School of E. E., University of Ulsan, Mugeo 2Dong, Namgu, Ulsan, Korea
*Next Generation Power Technology Center, Myongji University, Namdong, Yongin, Korea

ABSTRACT

This paper addresses a novel data error detection method in a substation. Unlike conventional methods, the proposed algorithm uses logical reasoning taking general structure of a substation into consideration.

A substation model is divided into several modules according to the voltage level and connection status. This modular approach enables us to utilize as many data as possible, because the data in main database do not contain all pieces of information measured at process buses. In addition, it reduces search space compared with centralized approach.

For the analog data correction, the conservative reasoning is adopted, which means that a data value is not evaluated to be "True" in case every evidence is proved. For digital data, a measure "Falseness" is suggested, by which a data can be evaluated to be false or not. Every member of a data set is ranked according to the Falseness which is calculated on every possible set of evidences. And the data with highest Falseness become candidates of bad data.

After the basic module is investigated, the original basic module is expanded by including one of adjacent modules. The new basic module is evaluated in the same manner as the original basic module. Simulations on a standard substation of KEPCO demonstrate the effectiveness of the proposed algorithm.

1. INTRODUCTION

The database of a Supervisory Control And Data Acquisition(SCADA) system or Energy Management System(EMS) contains numerous kinds of data such as bus voltages, line currents, switch status and others. If the database has any incorrect data, it may bring about dangerous or disastrous situations. The reason for the incorrect data is manifold. For example, there may happen some faults or errors in measuring, transferring or processing data. Late or no-update of changes in switch status can be another reason.

There exist a lot of interesting researches on error detection and correction of SCADA system or EMS data, most of which are focused on topological error handling in transmission systems. Some used state estimation or other analytic methods[1-3], and others, rule based techniques[4].

For more detailed and analytic approaches, switch and substation modeling methods are also presented[5-8]. Besides, there have been so many valuable research results presented. However, there exists little number of papers about BDP inside a substation. Considering recent tendency toward digitalization of substations, an effective method is required to preserve the soundness and accuracy of all the data in a substation.

This paper is about an algorithm of Bad Data Processing, abbreviated to BDP, that is going to be installed in a next generation digital substation. A methodology of BDP algorithm for analog and digital data is suggested, in this paper. Herein, The relations between one data and another are logically analyzed and the resultant error-symptoms relations are collected to form a rule base. A module bounded by measuring points is proposed for simple and systematic evaluation. The suggested algorithm is applied, through simulation, to a typical distribution substation working in Korea to show its effectiveness.

This paper is composed of 6 sections. The basic concept and philosophy is presented in section 2, followed by error analysis and rule base construction. In section 4, the whole procedure of the suggested algorithm and, in section 5, the application results to a substation is given. Conclusions and future works are summarized in the last section.

2. CONCEPT OF THE PROPOSED BAD DATA PROCESSING

This section is aimed at brief description of the basic ideas of the proposed Bad Data Processing(BDP). The main ideas in the proposed scheme can be summarized as 'conservative reasoning' and 'modular approach', which shall be explained in more detail in the following two subsections.

2.1. Conservative Reasoning

Essentially, BDP is a process to evaluate the truth of a data and to find out or correct data which are decided to be not true including unknown. In this, two aspects are possible. One is that every data is initially regarded as 'true' unless it is not detected to be 'false'. On the other hand, in the other aspect, only a data with authentic evidence is considered to be 'true'. We can define the former as affirmative reasoning and the latter as conservative one. In the affirmative reasoning, the number of proofs determines the efficiency, in other words, if we have small number of rules, the number of reasoning is also small. However, in case of even slight changes in rules or environment, a data in 'true' state should be evaluated again. Moreover, there is strong possibility, without sufficient information, that we can consider an 'actually false data' as a true data.

In the conservative reasoning, the initial state of a data is 'unknown' or 'false'. We should evaluate it using as many evidences as possible and then, we can assign the truth value of a data. Therefore, it seems that conservative reasoning requires so much time and effort to change the state of a data from 'unknown' or 'false' to 'true'. On the other hand, however, once a data is proved to be 'true', it can be trusted. It doesn't have to be evaluated again even if changes occur in rules or environment. This feature can reduce the dimension of a search space when many data are evaluated to be 'true' in the initial evaluation.

In the BDP problem of a substation, since we have only a small number of rules compared with the number of data, the affirmative reasoning may be thought to be advantageous. However, in reality, the conservative reasoning reduces the search space rapidly, because a large number of data are decided to be 'true'.

In this paper, after intensive case studies under the various conditions, the conservative reasoning is adopted. And in order to improve the efficiency of reasoning, a novel type of a 'modular approach' is suggested as explained in the next subsection.

2.2. Modular Approach

Since a substation consists of a lot of switches and elements, it is much convenient to divide it into some groups of elements, which is defined, in this paper, as "modules". There is an important point that it may not make sense to divide a substation into modules according to physical elements, e.g., transformers, primary bus, feeders, etc. It is because the measuring points are not coincident with the boundaries of the elements. Instead, we used the measuring points as the boundary of a module.

For convenience, we first define a 'Basic Module', which is selected among the 'highest' measuring points. The expression 'high' and 'low' mean, in this paper, 'to the source direction' and 'to the load direction', respectively. And then, find next measuring points along the connected elements to lower side or to the same level. In a typical distribution substation in Korea, the highest measuring points are located at the entrance of 154[kV] transmission lines to the substation. It should be emphasized that all the elements of a module should be connected electrically.

After a basic module is defined, next measuring points are searched. If a new measuring point is found, the elements between the basic module and this new measuring point make up an 'expanded module' along with the basic module.

In this way, we can construct a hierarchy of modules containing all the elements and measuring points in a substation. The proposed modular approach has an advantage of reducing the size of search space. First, we perform BDP algorithm with the data in the basic module, not in the whole substation. Then, only a small numbers of data remains 'unknown' or 'flase' state. If the module is expanded, only these 'unknown' and 'bad' data participate in the BDP algorithm together with the new data. Therefore, even if the total number of data increases, as we expand the basic module, the search space does not increase accordingly. The conservative reasoning explained in the previous subsection allows us to utilize the advantage of modular approach because we do not have to deal with the data decided to be 'true'.

2.3. Construction of Rule Base

When there exists an error in the database, even if we cannot point out the exact erroneous data, or correct value, we can find the fact that there is an error by consistency analysis of a group of data. And if we categorize and formulate the analysis results in a variety of error cases, we can construct a relation table between errors and the symptoms arisen from the errors, some of which are tabulated in table I. In this table, *V*, *I*, *P* and *Q* represent voltage, current, active and reactive power, respectively.

The rules used in the proposed algorithm are not based on human experts' knowledge, but based on this error-symptoms relations. One simple example is that;

Table 1. Examples of error-symptoms relations.

Type of error	Symptoms
V measurement error	There exist different *V* values between adjacent modules (or in a module) & Sum of *P* or *Q* has a nonzero value in a module
I measurement error	Sum of *I* has a nonzero value in a module & Sum of *P* or *Q* has a nonzero value in a module
V communication error	There exist different *V* values in a module
I communication error	Sum of *I* has a nonzero value in a module
Q communication error	Sum of *Q* has a nonzero value in a module
P communication error	Sum of *P* has a nonzero value in a module

'If an error occurs in current measurement at a point in a module, then the sum of currents flowing into and going out of the module is not identically zero.'

In real application, for conservative reasoning, the contraposition of the rule should be used, such as;

'If the sum of the currents flowing into and going out of a module is identically zero, then the current data of that module are <u>not false</u>.'

Herein, we did not define true but <u>not false</u> which is an intermediate state. It is because this is not the only evidence for all the current data to be true. Therefore, we cannot be convinced that a data is true until all the evidences reveal imply that the data is true.

The rules used in this paper are categorized into several groups according to their functions to form a more effective rule base as follows;

i) Rules for topological understanding
ii) Rules for detecting connection state
iii) Rules for measurement error detection
iv) Rules for communication error detection
v) Rules related to module expansion
vi) Rules for report generation

Since rules in group i), ii) and vi) are not new ones, any existing method can be utilized. Group iii) and iv) can be further divided into several subgroups, rules for voltage, rules for current, rules for switches and breakers, etc. In Basic Module expansion, some quantities should be calculated, for example, total current of the expanded part. The rules in group v) are utilized in manipulating such new pieces of information.

It should be noticed that, in some cases, we may find a group of data with errors, but not a single one. Even though the number of elements in the erroneous data group lessens with further reasoning, it may happen that we cannot reduce the number any more after the last reasoning. In this case, the proposed algorithm gives only a group of bad data candidates. For example, suppose there occurs a voltage measurement error at a bus. If we have sufficient voltage data of other modules, we can easily find the error and then correct it. However, in case we have only two voltage values with different magnitude, we cannot decide which is correct one. We just know that there must happen a problem in voltage measurement.

3. BAD DATA PROCESSING

As mentioned in the previous section, there are two types of data, analog and digital. In evaluating the analog data, it is prerequisite that the exact information on the connection states of elements should be given. And in digital data evaluation, the analog data should be utilized as the bases of estimation. This means that we cannot evaluate both of the types of data simultaneously. Therefore, a iterative method is used in the proposed algorithm.

Analog data are processed under the assumption that all the digital data are true. And then, with the results of analog data processing, digital data are evaluated. If there happens a correction of erroneous data, analog data are evaluated again, and so on. Even though it was almost impossible to find a logical proof, we found that 3 or 4 times of iteration were sufficient in the BDP of a module of a typical distribution substation.

3.1. Analog Data Processing(ADP)

In analog data processing, the sequence of rule application is very important. When we are to evaluate one data, we should use others. In this, if they have not been proved to be true, the evaluation itself may have no meaning. Simple but very useful fundamental law is to start from the easiest point.

Voltages are the easiest variables to be evaluated because not only their actual number but also the number of rules is small. Therefore, we start from the communication error check of each voltage in the basic module. It can be easily checked by comparing transmitted voltage values in the base module. If there is no error, then voltage measure check is the next step followed by evaluation of current values, P and Q. If all the measured values are evaluated, the module is expanded by adding another module to the basic module. Then, we repeat the procedure until all the data in the expanded module are evaluated. Of course, as mentioned before, the data that have been proved as good data in the basic module evaluation are not evaluated again, but are utilized as information. This procedure is repeated until all the data in a substation are evaluated. Then, the data which still remain in bad or unknown state are the candidates of bad data.

3.2. Digital Data Processing(DDP)

A digital data can be evaluated by comparing some kinds of real analog data with estimated ones. The estimated data are obtained by simple calculation assuming all the data are true. In case there happens a contradiction between real and estimated analog data, we can guess that the state of a switch (or of a group of switches) has an error by the rules explained in 2.3. Then, we change the state of that switch and calculate the analog data again. If the re-estimated analog data is the same as the measured data, we can conclude that the original state of that switch was ‘false’. Since not every switch has sufficient pieces of information to be used in evaluation, we can only evaluate a group of switches which contribute in constructing a module. For example, a main transformer has two circuit breakers, one is located in the primary and the other in the secondary side. If we find that there is an error in the states of those CBs, we cannot identify which one is wrong. We can only guess that an error happened in the state of one of them or both.

3.3. Combination of ADP and DDP

As explained in the previous subsections, analog data is processed under the assumption that all the digital data are correct, vice versa. Therefore, after one DDP, we have to re-confirm the other, since a DDP can change some data. For example, after a DDP, there may happen some changes in the data of the states of some switches, which causes some changes in the information of the electric connections. Therefore, we must perform ADP again under the new connection states.

The suggested algorithm utilizes iterative method of ADP and DDP, module by module. ADP and DDP are performed, in the first basic module, repeatedly until no data error is found. Then, an expanded basic module is constructed and the ADP and DDP iteration is performed in the same manner as in the first basic module. We can expand the basic module until it contains all the elements of a substation.

4. EXAMPLES

In this section, the application examples are given to show how the proposed algorithm works.

4.1. Description of the target substation

The suggested BDP algorithm is applied to a typical distribution substation model shown in figure 1.

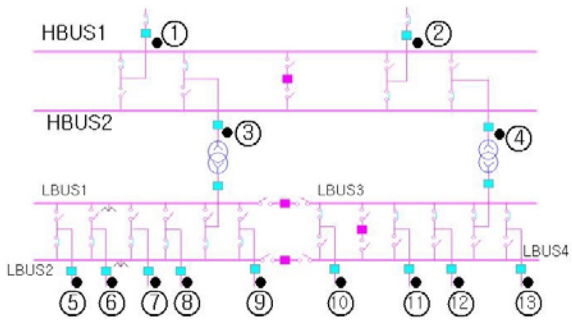

Figure 1. Simplified Diagram of the target substation.

This figure is a simplified diagram of a real substation in Seoul, Korea. In this figure, solid small circles are the measuring points, at which we can obtain the data of voltage, current, active and reactive power, and the figures in circles are the identification numbers of measuring points given for convenience. Actually, as mentioned before, the voltages are measured at the buses, HBUS 1 and LBUS2, LBUS4, in the given switch conditions. For the spatial problem, detailed data structure and topology identifying method used in this study are omitted.

4.2. ADP application results

For the spatial limitations, the ADP application is to be explained in detail. It is assumed that there happen two measurement errors, current at point 3 and voltage at LBUS2. The proposed algorithm starts with construction of the initial basic module.

i) **Construction of Basic Module(BM)**
The initial BM consists of point 1 to 4, as definition.

ii) **Initial Evaluation**
We can find that all the voltages are good, as assumed, on the other hand, since the sum of currents in the basic module is not zero, the currents are assigned to be false. Because the currents are false, every P and every Q are left to be in the unknown state.

iii) Expansion of BM
According to the topological data, initial BM is expanded to include points 5 to 9 through point 3.

iv) Evaluation for the new BM
In this case, the link point of the BM is point 3. As assumed, the voltage at point 3 is not equal to those of newly included points. However, since the voltage at point 3 has been already known to be a true data, we can conclude that the voltages of the feeders are false. Therefore, they are corrected and their states are changed into true. However, since the sum of currents of added part is not equal to the current flowing though point3, in this example, all the current states are still assigned to be false. Consequently, every P and every Q are still in the unknown states.

v) Expansion of BM
The BM is expanded to include points 10 to 13 through point 4.

iv) **Evaluation for the new BM**
After voltage check, all the voltages are assigned to be false. Current check provides a good piece of information. We can find that the current flowing through point 4 is identical to the sum of point 10 to 13, which implies that all the current values participated in this calculation are good. Then, the state of current at point 4 as well as those at point 10 to 13 is assigned true. Besides, by summing of the currents in the new BM, we can conclude that all the currents at the boundary measuring points are true. Finally, it is revealed that the current measurement at point 3 was false.

Table 2. Results of every step.

Step	Measuring Points Included in BM	Data in the 'good' state	Data in the 'bad' state	Data in the 'unknown' state
i)	1,2,3,4			
ii)	1,2,3,4	*V* at 1,2,3,4	*I* at 1,2,3,4	*P&Q* at 1,2,3,4
iii)	1,2,3,4, 5, 6, 7, 8, 9			
iv)	1,2,3,4, 5, 6, 7, 8, 9	*V* at 1,2,3,4 5,6,7,8,9	*I* at 1,2,3,4 5,6,7,8,9	*P&Q* at 1,2,3,4 5,6,7,8,9
v)	1,2,3,4, 5, 6, 7, 8, 9 10,11,12,13	***All anlaog data***		

As the last procedure, Every P and every Q are evaluated and assigned to be true.

The results of every step are tabulated in table 2. With excessive application to various kinds of errors, the suggested algorithm is proved to be useful for error detection and data correction. Other application examples show similar results including DDP and combination of ADP and DDP, which are omitted for spatial problem.

5. CONCLUSIONS

In this paper, a Bad Data Processing(BDP) algorithm is suggested for substations. Rules for detection and correction of bad data are derived through logical error-symptoms analysis. These rules are applied in a conservative manner so that the search space may lessen as the proposed step-wise reasoning algorithm goes on. A module separated by measuring points is also suggested for systematic evaluation.

The suggested algorithm was proved to be useful in both convenience and accuracy from excessive simulation results in a typical substation in Korea.

ACKNOWLEDGEMENT

Authors would like to thank to Korea Ministry of Science and Technology and Korean Science and Engineering Foundation for their support through ERC program

.

REFERENCES

1. Felix F. Wu and Wen H. E. Liu, "Detection of topologic errors by state estimation", *IEEE Trans. Power Systems*, vol. 4, pp. 176-183, Feb., 1989.
2. Wen-Hsiung E, Liu, Felix F. Wu and Shau M. Lun, "Estimation of parameter errors from measurement residuals in state estimation", *IEEE Trans. on Power Systems*, vol. 7, pp. 81-89, Feb., 1992.
3. Antonio.S. Costa and Jose A. Leao, "Identification of topology errors in power system state estimation", *IEEE Trans. Power Systems*, vol. 8, pp. 1531-1538, Nov., 1993.
4. N. Singh and H. Glavitsch, "Detection and identification of topological errors in online power system analysis", *IEEE Trans. Power Systems*, vol. 6, pp.324-331, Feb., 1991.
5. A. Monticelli and A. Garcia, "Modeling zero impedance branches in power system state estimation", *IEEE Trans. Power Systems*, vol. 6, pp.1561-1570, Nov. 1991.
6. A. Monticelli, "Modeling circuit breakers in weighted least square state estimation", *IEEE Trans. Power Systems*, vol. 8, 1143-1149, Aug. 1993.
7. Antonio G. Exposito, Antonio V. Jaen, "Reduced substation models for generalized state estimation", *IEEE Trans. Power Systems*, vol. 16, pp. 839-846, Nov. 2001.
8. Antonio V. Jaen and Antonio G. Exposito, "Implicitly constrained substation model for state estimation", *IEEE Trans. Powr Systems*, vol. 17, pp.850-856, Aug. 2002.