2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Cyber Attack Detection and Faults Diagnosis in Power Networks by Using State Fault Diagnosis Matrix

Yuki Fujita[1], Toru Namerikawa[1] and Kenko Uchida[2]

Abstract— **This paper discusses cyber attacks and faults diagnosis method for networked electrical power systems. In this paper, a power system is divided into some areas, and then, some parameters are estimated by Kalman filter in each area. Then we propose a diagnosis method by using called state and output fault diagnosis matrices which are composed of the estimated values, output matrix and so on. These diagnosis matrices can immediately judge which nodes of which area is attacked. Finally, some simulation results are shown to validate the proposed approach.**

I. INTRODUCTION

The power network system is one of the most complex and largest machine engineered by humankind in the 20th century. If troubles occur in this huge, complex system, it may result in large power failure and the loss is inscrutable. Therefore, the power grid is a system that should value safety most. And to avoid the power problems some dangerous factors should be studied and effective methods should be developed. In recent years, especially, the risk of cyber attacks that can cause information falsifications and abuse have been increasing on the background of the developments of information technology. And some country should consider the massive breakdown of the infrastructure because of a natural disaster such as earthquake, typhoon, thunder, and so on. Moreover, the next generation power grid using natural energy and advanced information-communication technology (such as smart grid) is expected in Japan because the operation rates of some nuclear plants are steadily declining [1],[2]. These power network system is known to have a lot of merits, for example, energy saving and blackout tolerance [3] because each customer generates some electric power by using photovoltaic power generator. However, these smart grid has more risks to be attacked than the existing system owing to underlying information-communication technology. Therefore, the ways to detect and diagnose the cyber attacks and faults are still much more expected to be developed and we discuss about this in this paper.

In order to detect and diagnose the cyber attacks for the power grid a lots of methods have been proposed, for example [4]. However, some paper discuss this problem in the case that cyber attacks is given only to the state, and some other papers is only about the output. From these background, a way to detect is proposed that covers both of the attacks [5]. But this method is centralized and needs all of the information about the generators and loads, and using all of the data may be difficult in a real situation. Therefore, the distributed methods (such as [6]) are expected to be developed for the next generation power grid because the system gradually gets more massive, complex and advanced. Moreover, the distributed detection is more expected than the centralized one from the view point of the fault tolerance. Therefore, a distributed method to detect and diagnose the cyber attacks promptly is expected even if the abnormal signals are given to the both of the state and output.

From these backgrounds, a distributed method to diagnose immediately which nodes is attacked is proposed. And we cope with the case that there are both cyber attacks for state and output. In this paper, at first, the power network system is divided into some areas and the state is estimated in each area by using Kalman filter. Then, the called fault diagnosis matrix is defined by combining the estimated parameter, output matrix and so on appropriately. This fault diagnosis matrix is defined for state and output individually. Therefore, we use two diagnosis matrices, and in this paper, we call them state fault diagnosis matrix and output fault diagnosis matrix [7] respectively. It becomes possible to diagnose individually the two kinds of abnormal attack signals by using these state fault diagnosis matrix and output fault diagnosis matrix. The key property of these diagnosis matrices is that the effects of the attack signals appear in the diagonal elements of these diagnosis matrices immediately if some evil signals are given to the power system. Therefore, in this method, the upper and lower bounds for the trace of these diagnosis matrices are set, and then, we give a decision based on whether the value of trace is in the range or not. Although the threshold of output fault diagnosis matrix is constant, that of state fault diagnosis matrix changes dynamically to cut down the time delay. The main merits of this method are that both of the attack signals (for state and output) can be diagnosed individually and this is a distributed method not centralized one.

II. PROBLEM STATEMENT

In a power network with N generators, we associate with generator j its inertia constant M_i, its electrical power output P_{ej}, its mechanical power input P_{mj}, its damping constant D_j, and its relative rotor angle δ_j measured with respect to a synchronously rotating reference frame with frequency f. These parameters are given in each power network and the synchronous frequency f is typically given as 50Hz or 60Hz. Then, the swing equation which is the rotor dynamics

*This work was supported by JST CREST

[1]Y. Fujita and T. Namerikawa are with Department of System Design Engineering, Keio University, Kanagawa, Japan namerikawa at nl.sd.keio.ac.jp

[2]K. Uchida is with Department of Electrical Engineering and Bioscience, Waseda University, Tokyo, Japan kuchida at waseda.jp

of generator j are given as

$$M_j\ddot{\delta}_j = P_{mj} - D_j\dot{\delta}_j - P_{ej} \tag{1}$$

$$P_{ej} = \sum_{k\in\mathcal{N}_j} \mid V_j \mid\mid V_k \mid \{G_{jk}\cos(\delta_j - \delta_k) + B_{jk}\sin(\delta_j - \delta_k)\} \tag{2}$$

Where all terms are in per unit value and $j \in \{1, ..., N\}$. And $\mathcal{N}_j$ means neighbor nodes of node j, $\mid V_j \mid$ is voltage of node j, G_{jk} and B_{jk} are mutual conductance and susceptance.

In this paper, we consider IEEE 118 bus power network system which consists of N generators and M loads although a huge number of generators and loads are generally connected to power network. Then, suppose that the power network system is divided into L areas. This power network system is shown in Figure 1.

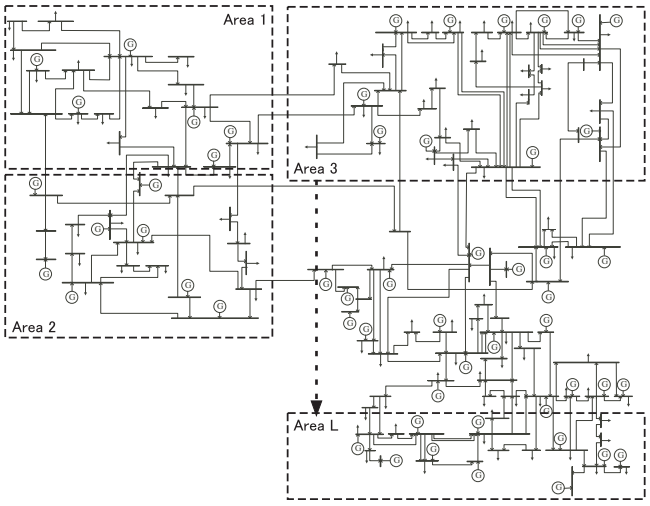

Fig. 1. Division of IEEE 118 bus power network system into L areas

Then, the next assumptions are set for the divided power network system.

Assumption 1:

iDThis power network system is lossless

iiDThe relative rotor angle differences between the neighboring nodes are sufficiently small.

iiiDnode $i(\in AreaL)$ can get the information of node $j(\in AreaL)$

where assumption 1-i is due to ultrahigh-voltage power transmission, so we can deal with the conductances which is included in the P_{ej} as zero. And the non-linear dynamics of generators and loads can be approximated as linear dynamics due to this assumption 1-ii. Therefore, the power output (2) can finally be rewritten as

$$P_{ej} = \sum_{k\in\mathcal{N}_j} \mid V_j \mid\mid V_k \mid B_{jk}(\delta_j - \delta_k) \tag{3}$$

In this paper, the dynamics of all nodes can be described as swing equation (1)-(2) because we consider that all nodes are general motor load in this power grid. Therefore, the swing equation in each node is given as

$$M_j\ddot{\delta}_j + D_j\dot{\delta}_j = P_{mj} + P_{in,j} - \sum_{k\in\mathcal{N}_j} \mid V_j \mid\mid V_k \mid B_{jk}(\delta_j - \delta_k) \tag{4}$$

where $j \in \{1, 2, ... , N+M\}$ and the $P_{in,j}$ is the power flow into node j from the nodes which belongs to the other areas. Therefore, if the node j is not connected to any other area's nodes, this $P_{in,j}$ is calculated as zero.

Then, dealing with the swing dynamics (4) in each area as linear time-fixed discrete dynamics system, the system of area i can be described as [8]

$$x^i_{k+1} = A^i x^i_k + B^i u^i_k + w^i_k + E^i_k h^i_k \tag{5}$$

$$y^i_k = C^i x^i_k + v^i_k + F^i_k g^i_k \tag{6}$$

where $i \in \{1, 2, ... , L\}$ means areas. In addition, Let $\mid i \mid$ be the numbers of the nodes which is in area i, $x^i \in \mathbb{R}^{2|i|}$, $A^i \in \mathbb{R}^{2|i|\times|2|i|}$, $B^i \in \mathbb{R}^{2|i|\times 2|i|}$, $u^i \in \mathbb{R}^{2|i|}$are defined as

$$x_j = \begin{bmatrix} \delta_j & \omega_j \end{bmatrix}^T, x^i = \begin{bmatrix} x_1^T & , & x_{|i|}^T \end{bmatrix}^T \tag{7}$$

$$A^i = (M^i)^{-1}\left\{D^i - \mathcal{W}^i\mathcal{L}^{i'}\right\} \tag{8}$$

$$m_j = \begin{bmatrix} 1 & 0 \\ 0 & M_j \end{bmatrix}, M^i = \begin{bmatrix} m_1 & & \\ & \ddots & \\ & & m_{|i|} \end{bmatrix} \tag{9}$$

$$d_j = \begin{bmatrix} 0 & 1 \\ 0 & -D_j \end{bmatrix}, D^i = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_{|i|} \end{bmatrix} \tag{10}$$

$$\mathcal{W}^i = \left\{\mathcal{K}^i \otimes \mathcal{G}\right\}\left\{(\mathcal{L}^i)^\dagger \otimes \mathcal{H}\right\} \tag{11}$$

$$\mathcal{K}^i_{(s,t)} = \begin{cases} \sum_{t\in\mathcal{N}_s} \mid V_s \mid\mid V_t \mid B_{st} & t = s \\ -\mid V_s \mid\mid V_t \mid B_{st} & t \in \mathcal{N}_s \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

$$\mathcal{F} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \mathcal{G} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \mathcal{H} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{13}$$

$$\mathcal{L}^{i'} = \mathcal{L}^i \otimes \mathcal{F}, \mathcal{L}^i_{(s,t)} = \begin{cases} \sum_{t\in\mathcal{N}_s} 1 & t = s \\ -1 & t \in \mathcal{N}_s \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

$$P_j = \begin{bmatrix} P_{in,j} & P_{mj} \end{bmatrix}^T, u^i = \begin{bmatrix} P_1^T & , & P_{|i|}^T \end{bmatrix}^T \tag{15}$$

$$B^i = M^{-1}B' \ , \ B' = diag[b, , , b], \ b = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \tag{16}$$

where $j \in \{1, ... , \mid i \mid\}$, w_j is angular frequency, $\mathcal{L}$ is graph laplacian, † means resolvent matrix [9], and output

matrix is unit matrix. And vector x^i is composed of relative rotor angle and its temporal differentiation, A^i is composed of inertia constants, damping constants, graph laplacian, and so on, u^i is input consists of mechanical power input and power flow form other area.

With respect to noise, $w_k^i \in \mathbb{R}^{2|i|}$ is process noise, and its covariance is $W^i \geq 0$, the average is zero. And this is white-noise depends on normal probability distribution. And $v_k^i \in \mathbb{R}^{2|i|}$ is observation noise, and its covariance is $V^i \geq 0$, the average is zero. And this is white-noise depends on normal probability distribution.

$E_k^i \in \mathbb{R}^{2|i|\times 2|i|}$, $h_k^i \in \mathbb{R}^{2|i|}$, $F_k^i \in \mathbb{R}^{2|i|\times 2|i|}$, $g_k^i \in \mathbb{R}^{2|i|}$ included in the system (5)-(6) are, respectively, the cyber attacks or faults signals for state and the cyber attacks or faults signals for output. These signals are not necessarily Gaussian white-noise. In this paper, we assume the cases such as state changes due to the information falsifications and abuse by the cyber attacks [10] or some bias on observation data due to rapid pace of situation changes. These cases can be discussed and dealt with by using these signal E_k^i, h_k^i, F_k^i, g_k^i. For example, if the attack signal for output is not zero suck as $F_k^i g_k^i > 0$ at step k, this means that at least one abnormal signal is given to sensor of some nodes at that step. Then, we set the next assumptions for the power network system modeled as equations (5)-(6).

Assumption 2:

iD$E[v_k^i v_s^{iT}] = E[w_k^i w_s^{iT}] = 0 \quad (k \neq s)$

iiD$E[v_k^i w_s^{iT}] = E[v_k^i h_s^{iT}] = E[v_k^i g_s^{iT}] = E[w_k^i h_s^{iT}]$
$= E[w_k^i g_s^{iT}] = 0$

iiiD$E[x_0 v_s^{iT}] = E[x_0 w_s^{iT}] = E[x_0 h_s^{iT}]$
$= E[x_0 g_s^{iT}] = 0$

ivDE$[\left\{h_k^i - E(h_k^i)\right\}\left\{h_k^i - E(h_k^i)\right\}^T] = H_k^i \geq 0$
E$[\left\{g_k^i - E(g_k^i)\right\}\left\{g_k^i - E(g_k^i)\right\}^T] = G_k^i \geq 0$

Assumption 3:

iD$(A, W^{\frac{1}{2}})$ is reachable
iiD(C, A) is detectable

Then, we define the diagnosis problem of cyber attacks and faults which is dealt with in this paper.

Problem 1:
Assume that the power network system model is described as equations (5)-(6) under the assumption 1-3, if at least one cyber attack or faults signal is given to the power gird (the cases that $E_k^i h_k^i \neq 0$ or $F_k^i g_k^i \neq 0$), define a diagnostic signal which can promptly identify which node of which area has gotten the abnormal signal.

III. PROPOSED APPROACH

In this section, we propose the method to diagnose the cyber attacks and fault signals. In this method, we define the state fault diagnosis matrix and output fault diagnosis matrix at first, and then, the diagnostic signal is defined by using these two matrices. In the process of the definition of the diagnostic signal, we refer to way to fix the threshold dynamically.

A. STATE FAULT DIAGNOSIS MATRIX

Let the $\hat{x}_{k|k-1}^i$, $P_{k|k-1}^i$, and K_k^i be, respectively, estimated state values of area i at step $k-1$, estimated covariance values of area i at step $k-1$, and Kalman gain of area i. Then, the estimate equations of Kalman filter at step k are described as

$$\hat{x}_{k+1|k}^i = A^i \hat{x}_{k|k}^i + B^i u_k^i \tag{17}$$

$$\hat{x}_{k|k}^i = \hat{x}_{k|k-1}^i + K_k^i \left\{ y_k^i - C_k^i \hat{x}_{k|k-1}^i \right\} \tag{18}$$

$$S_k^i = cov(y_k^i - C_k^i \hat{x}_{k|k-1}^i) \tag{19}$$

$$K_k^i = P_{k|k-1}^i C_k^{iT} \left\{ S_k^i \right\}^{-1} \tag{20}$$

$$P_{k+1|k}^i = A^i P_{k|k}^i A^{iT} + W_k^i \tag{21}$$

$$P_{k|k}^i = P_{k|k-1}^i - K_k^i C_k^i P_{k|k-1}^i \tag{22}$$

Then, covariance is a worth noting factor to consider the relationship between abnormal signal $E^i h^i$ and estimation error. Let $\eta_k^i = x_k^i - \hat{x}_{k|k-1}^i$, $\eta^{i'}{}_k = x_k^i - \hat{x}_{k|k}^i$ be estimation errors, and then, the estimation error η_k^i can be rewritten as

$$\begin{aligned}\eta_k^i &= x_k^i - \hat{x}_{k|k-1}^i \\ &= A(x_{k-1}^i - \hat{x}_{k-1|k-1}^i) + w_{k-1}^i + E_{k-1}^i h_{k-1}^i \\ &= A\eta_{k-1}' + w_{k-1}^i + E_{k-1}^i h_{k-1}^i \end{aligned} \tag{23}$$

This equation shows that how does the abnormal signal $E_{k-1}^i h_{k-1}^i$ effect on the estimation error. And then, by using equation (23) the covariance $P_{k|k-1}^i$ can be recalculated as

$$\begin{aligned}P_{k|k-1}^i &= cov(\eta_k^i) \\ &= A^i P_{k-1|k-1}^i A^T + W_{k-1}^i + E_{k-1}^i H_{k-1}^i E_{k-1}^i \\ &+E\left[A^i \eta_{k-1}^{i'} \left(h_{k-1}^{iT} E_{k-1}^{iT} - E\left[h_{k-1}^{iT} E_{k-1}^{iT}\right]\right)\right. \\ &\left.-E\left[A^i \eta_{k-1}^{i'}\right] \left(h_{k-1}^{iT} E_{k-1}^{iT} - E\left[h_{k-1}^{iT} E_{k-1}^{iT}\right]\right)\right] \\ &+E\left[E_{k-1}'^i h_{k-1}^i \left(\eta_{k-1}^{i'T} A^{iT} - E\left[\eta_{k-1}^{i'T} A^{iT}\right]\right)\right. \\ &\left.-E\left[E_{k-1}'^i h_{k-1}^i\right] \left(\eta_{k-1}^{i'T} A^{iT} - E\left[\eta_{k-1}^{i'T} A^{iT}\right]\right)\right] \end{aligned} \tag{24}$$

Then, we define the state fault diagnosis matrix of area i at step k as N_k^i by using the covariance $P_{k|k-1}^i$.

State Fault Diagnosis Matrix :

$$\begin{aligned}N_k^i &:= P_{k|k-1}^i - A^i P_{k-1|k-1}^i A^{iT} - \hat{W}_{k-1}^i \\ &= W_{k-1}^i - \hat{W}_{k-1}^i + E_{k-1}^i H_{k-1}^i E_{k-1}^i \\ &+E\left[A^i \eta_{k-1}^{i'} \left(h_{k-1}^{iT} E_{k-1}^{iT} - E\left[h_{k-1}^{iT} E_{k-1}^{iT}\right]\right)\right. \\ &\left.-E\left[A^i \eta_{k-1}^{i'}\right] \left(h_{k-1}^{iT} E_{k-1}^{iT} - E\left[h_{k-1}^{iT} E_{k-1}^{iT}\right]\right)\right] \end{aligned}$$

$$+E\left[E'^{i}_{k-1}h^{i}_{k-1}\left(\eta^{i'T}_{k-1}A^{iT}-E\left[\eta^{i'T}_{k-1}A^{iT}\right]\right)\right.$$
$$\left.-E\left[E'^{i}_{k-1}h^{i}_{k-1}\right]\left(\eta^{i'T}_{k-1}A^{iT}-E\left[\eta^{i'T}_{k-1}A^{iT}\right]\right)\right] \quad (25)$$

where $\hat{W}^{i}_{k-1}$ is covariance of estimated process noise that is white-noise and the average is zero. that is $\hat{W}^{i}_{k-1} = E\left[\hat{w}^{i}_{k}\hat{w}^{iT}_{k}\right]$. From above equation, it can be confirmed that the state fault diagnosis matrix N^{i}_{k} is composed of covariance of abnormal signals and correlation between estimation error and abnormal signal except for the covariance of process noise and its estimation. Therefore, if a cyber attacks or fault signal occurs in some nodes, the effects of this abnormal signal $E^{i}_{k}h^{i}_{k}$ are directly reflected in the correspondent matrix element.

Property 1:
Suppose that the state x^{i}_{k} and the estimated value of the state $\hat{x}^{i}_{k|k}$ have no correlation with the cyber attacks and fault signals $E^{i}_{k}h^{i}_{k}$ under the assumption 2 and 3. In this case, the state fault diagnosis matrix increases as the abnormal signals increase.

$$E^{1}_{k}H^{1}_{k}E^{1T}_{k} \leq E^{2}_{k}H^{2}_{k}E^{2T}_{k} \longrightarrow N^{1}_{k} \leq N^{2}_{k} \quad (26)$$

Proof 1:
If the state x^{i}_{k} and the estimated value of that $\hat{x}^{i}_{k|k}$ have no correlation with the abnormal signals $E^{i}_{k}h^{i}_{k}$, the state fault diagnosis matrix (25) can be simplified by recalculation.

$$N^{i}_{k} = W^{i}_{k-1} - \hat{W}^{i}_{k-1} + E^{i}_{k-1}H^{i}_{k-1}E^{i}_{k-1} \quad (27)$$

Therefore, the magnitude relationship between state fault diagnosis matrix N^{i}_{k} and abnormal signal $E^{i}_{k-1}H^{i}_{k-1}E^{i}_{k-1}$ corresponds. In addition, the two noises of W^{i}_{k-1} and its estimation $\hat{W}^{i}_{k-1}$ cancel each other out.

B. OUTPUT FAULT DIAGNOSIS MATRIX

The estimate equation of Kalman filter is almost same as in the previous section. But the covariance of the observation error S^{i}_{k} is slightly different in the case that cyber attacks and fault signals are given to the system. In these case, the S^{i}_{k} is described as

$$\begin{aligned}
S^{i}_{k} &= cov(y^{i}_{k} - C^{i}_{k}\hat{x}^{i}_{k|k-1}) \\
&= C^{i}P^{i}_{k|k-1}C^{iT} + V^{i}_{k} + F^{i}_{k}G^{i}_{k}F^{iT}_{k} \\
&= +E\left[C^{i}\eta^{i}_{k}\left(g^{iT}_{k}F^{iT}_{k} - E\left[g^{iT}_{k}F^{iT}_{k}\right]\right)\right. \\
&\quad \left.-E\left[C^{i}\eta^{i}_{k}\right]\left(g^{iT}_{k}F^{iT}_{k} - E\left[g^{iT}_{k}F^{iT}_{k}\right]\right)\right] \\
&\quad +E\left[g^{i}_{k}F^{i}_{k}\left(\eta^{iT}_{k}C^{iT}_{k} - E\left[\eta^{iT}_{k}C^{iT}_{k}\right]\right)\right. \\
&\quad \left.-E\left[g^{i}_{k}F^{i}_{k}\right]\left(\eta^{iT}_{k}C^{iT}_{k} - E\left[\eta^{iT}_{k}C^{iT}_{k}\right]\right)\right] \quad (28)
\end{aligned}$$

Then, we define the output fault diagnosis matrix of area i at step k as M^{i}_{k} by using the covariance of observation error S^{i}_{k}.

Output Fault Diagnosis Matrix :

$$\begin{aligned}
M^{i}_{k} &:= S^{i}_{k} - C^{i}_{k}P^{i}_{k|k-1}C^{iT}_{k} - \hat{V}^{i}_{k} \\
&= V^{i}_{k} - \hat{V}^{i}_{k} + F^{i}_{k}G^{i}_{k}F^{iT}_{k} \\
&\quad +E\left[C^{i}_{k}\eta^{i}\left(g^{iT}_{k}F^{iT}_{k} - E\left[g^{iT}_{k}F^{iT}_{k}\right]\right)\right. \\
&\quad \left.-E\left[C^{i}_{k}\eta^{i}\right]\left(g^{iT}_{k}F^{iT}_{k} - E\left[g^{iT}_{k}F^{iT}_{k}\right]\right)\right] \\
&\quad +E\left[g^{i}_{k}F^{i}_{k}\left(\eta^{iT}C^{iT}_{k} - E\left[\eta^{iT}C^{iT}_{k}\right]\right)\right. \\
&\quad \left.-E\left[g^{i}_{k}F^{i}_{k}\right]\left(\eta^{iT}C^{iT}_{k} - E\left[\eta^{iT}C^{iT}_{k}\right]\right)\right] \quad (29)
\end{aligned}$$

where $\hat{V}^{i}_{k-1}$ is covariance of estimated observation noise that is white-noise and the average is zero. that is $\hat{V}^{i}_{k-1} = E\left[\hat{v}^{i}_{k}\hat{v}^{iT}_{k}\right]$. From above equation, it can be confirmed that the output fault diagnosis matrix M^{i}_{k} is composed of covariance of abnormal signals and correlation between estimation error and abnormal signal except for the covariance of observation noise and its estimation. Therefore, if a cyber attacks or fault signal occurs in some nodes, the effects of this abnormal signal $F^{i}_{k}g^{i}_{k}$ are directly reflected in the correspondent matrix element.

Property 2:
Suppose that the state x^{i}_{k} and the estimated value of the state $\hat{x}^{i}_{k|k}$ have no correlation with the cyber attacks and fault signals $F^{i}_{k}g^{i}_{k}$ under the assumption 2 and 3. In this case, the state fault diagnosis matrix increases as the abnormal signals increase.

$$F^{1}_{k}G^{1}_{k}F^{1T}_{k} \leq F^{2}_{k}G^{2}_{k}F^{2T}_{k} \longrightarrow M^{1}_{k} \leq M^{2}_{k} \quad (30)$$

Proof 2:
If the state x^{i}_{k} and the estimated value of the state $\hat{x}^{i}_{k|k}$ have no correlation with the abnormal signals $F^{i}_{k}g^{i}_{k}$, the output fault diagnosis matrix (29) can be simplified by recalculation.

$$M^{i}_{k} = V^{i}_{k-1} - \hat{V}^{i}_{k-1} + F^{i}_{k-1}G^{i}_{k-1}F^{i}_{k-1} \quad (31)$$

Therefore, the magnitude relationship between state fault diagnosis matrix M^{i}_{k} and abnormal signal $F^{i}_{k-1}G^{i}_{k-1}F^{i}_{k-1}$ corresponds. In addition, the two noises of V^{i}_{k-1} and its estimation $\hat{V}^{i}_{k-1}$ cancel each other out.

C. DIAGNOSTIC SIGNAL AND THRESHOLD

First, we define the state and output fault diagnosis matrix for node j as N^{i_j}, M^{i_j}, respectively. Although the matrices N^{i}, M^{i} can reflect the system trouble of area i, but these two matrix can not diagnose which node of the area i is attacked. Therefore, in order to diagnose each node not each area, we redefine N^{i_j}, M^{i_j} as the state and output fault diagnose matrix, respectively. These matrices are given as

$$N^{i_j} = \begin{bmatrix} N^{i}_{(2j-1,2j-1)} & N^{i}_{(2j-1,2j)} \\ N^{i}_{(2j,2j-1)} & N^{i}_{(2j,2j)} \end{bmatrix}$$

$$M^{i_j} = \begin{bmatrix} M^i_{(2j-1,2j-1)} & M^i_{(2j-1,2j)} \\ M^i_{(2j,2j-1)} & M^i_{(2j,2j)} \end{bmatrix}$$

where the ^{i_j} means the jth node of area i, and $(\ ,\ )$ means the matrix element.

Then, we can define the diagnostic signal R^j_k as the solution for *Problem 1* via these two fault diagnosis matrices.

$$R^j_k := \begin{cases} 0 & if \quad \mathcal{M}^{\min} \le traceM^{i_j}_k \le \mathcal{M}^{\max} \\ & and \quad \mathcal{N}^{\min}_k \le traceN^{i_j}_k \le \mathcal{N}^{\max}_k \\ 1 & otherwise \end{cases} \tag{32}$$

where the threshold of output fault diagnosis matrix $\mathcal{M}^{\min}, \mathcal{M}^{\max}$ is the constant parameter designed in advance. Therefore, the sensitivity of this diagnosis is decided by these two designed parameters. If the threshold was set strictly, almost all of the abnormal signals would be diagnosed even if the signals were so small, but sometimes, a noise might be identified as a kind of cyber attack. And if the threshold was set slackly, the converse would happen.

On the other hand, the threshold of the state fault diagnosis matrix $\mathcal{N}^{\min}, \mathcal{N}^{\max}$ is decided in each step referring to the changes of covariance P^i because cyber attacks and fault signals have some effects on estimation accuracy. Therefore, the threshold in next step is given as

$$\mathcal{N}^{\max}_{k+1} = \begin{cases} \mathcal{N}^{\max}_k - \mathcal{N} & if \ P_{k+1|k} < \mathcal{C}1 \quad and \\ & \mid P_{k+1|k} - P_{k|k-1} \mid < \mathcal{C}2 \\ \mathcal{N}^{\max}_k + \mathcal{N} & otherwise \end{cases} \tag{33}$$

$$\mathcal{N}^{\min}_{k+1} = \begin{cases} \mathcal{N}^{\min}_k + \mathcal{N} & if \ P_{k+1|k} < \mathcal{C}1 \quad and \\ & \mid P_{k+1|k} - P_{k|k-1} \mid < \mathcal{C}2 \\ \mathcal{N}^{\min}_k - \mathcal{N} & otherwise \end{cases} \tag{34}$$

where $\mathcal{N}, \mathcal{C}1, \mathcal{C}2$ are the designed parameters. The $\mathcal{N}$ are the parameters which decide how much next step's threshold will change compared to current one and all of these $\mathcal{N}$ are not necessarily same. And the $\mathcal{C}1, \mathcal{C}2$ are the parameter which judge if the threshold is able to be change to be strict or not. That is, if the covariance $P_{k+1|k}$ is less than $\mathcal{C}1$ and the difference of the covariance $P_{k+1|k}$ and $P_{k|k-1}$ is less than $\mathcal{C}2$, this means that estimation is accurate and that's why this means that there are no cyber attacks or faults in this power network model and there no problems if the threshold can be changed to be narrow.

IV. SIMULATION RESULTS

Suppose that the power network system is composed of 12 buses and divided into 4 areas as in Figure 2.

In these simulations, the upper bound of the output fault diagnosis matrix $\mathcal{M}^{\max}$ is 0.5, the lower bound $\mathcal{M}^{\min}$ is -0.5, the initial upper bound of the state fault diagnosis matrix $\mathcal{N}^{\max}_0$ is 0.5 and the initial lower bound $\mathcal{N}^{\min}_0$ is -0.5. Then, there are 4 $\mathcal{N}$ in the equations (33),(34) and these parameters are designed from 0.005 to 1.0 in this section, $\mathcal{C}1$ and $\mathcal{C}2$ are 0.005 and 0.3 respectively. And finally, sampling time is set as $0.1[s]$.

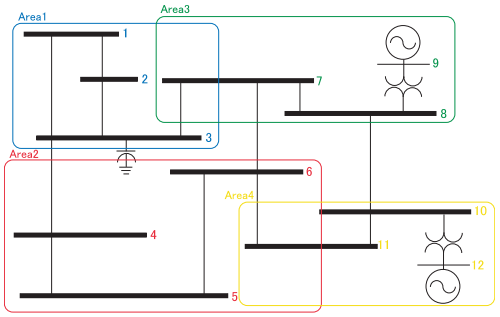

Fig. 2. Division of 12 bus power network system into 4 areas

A. RESULTS 1

In this section, we show some simulation results when the attacks and fault signals are given only to the state. the case of the output is shown in the next section.

The attack signals for states Eh are given to each nodes and these are shown in Figure 3. This figure shows that the abnormal signal is given only to node 1.

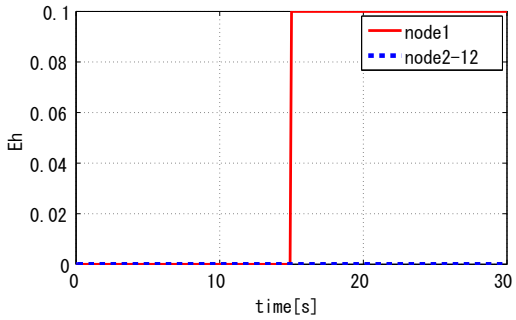

Fig. 3. Attack or fault signals for state

The result of the state fault diagnosis matrices and the thresholds is shown in Figure 4 and the diagnostic signals are shown in Figure 5.

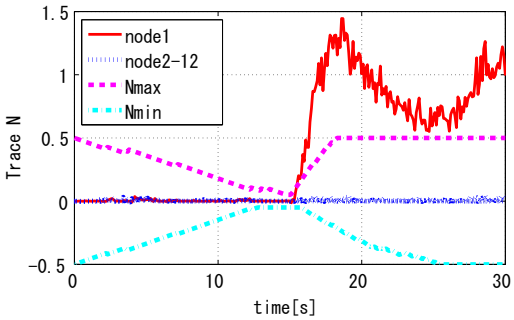

Fig. 4. Traces of state fault diagnosis matrix and thresholds

In the Figure 4, the red line is the trace of the state fault diagnosis matrix of node 1 and the blue broken lines show the those of the other nodes. And the other two broken lines show the shift of threshold. From this figure, it can be seen that $\mathcal{N}^{1_1}$ (described by red line) is the only diagnosis matrix

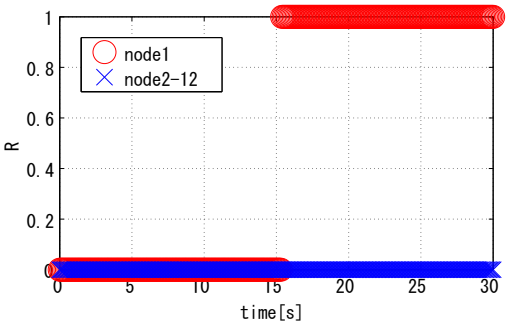

Fig. 5. Diagnostic signals

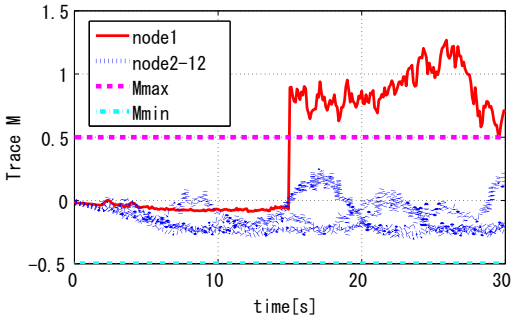

Fig. 7. Traces of output fault diagnosis matrix and thresholds

that is increasing right after Eh is added. And it can be also confirmed that the threshold changes depending on the estimation accuracy because that continues to narrow until Eh is given. Comparing above two figures, we can state that the signal R can diagnose the attack Eh and the dynamical changes of threshold get the diagnosis quickly.

B. RESULTS 2

In this section, we show some simulation results when the attacks and fault signals are given only to the output. The attack signals for outputs Fg are given to each nodes and these are shown in Figure 6. This figure shows that the abnormal signal is given only to node 1. And the results of the output fault diagnosis matrices and the thresholds are shown in Figure 7 and the diagnostic signals are shown in Figure 8.

From these figures, it can be seen that the diagnostic signal R is able to identify which nodes of which areas is not normal and the threshold does not have to change in case of the Fg because the output fault diagnosis matrix is increasing so quickly right after the Fg is given.

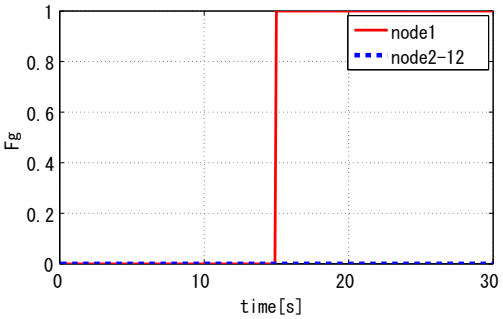

Fig. 6. Attack or fault signals for state

V. CONCLUSIONS

This paper presents a way to detect and diagnose the abnormal signals from outside of the power network system such as cyber attacks. And this method can also work in the case of the mechanical faults. This proposed method uses state and output fault diagnosis matrices which are composed of the estimated value by Kalman filter. The feature of this

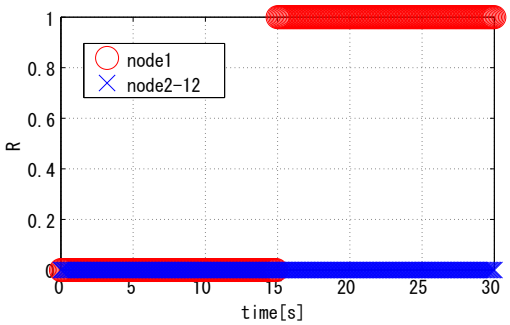

Fig. 8. Diagnostic signals

proposal is that the cyber attacks are individually detected and diagnosed in each area. And this distributed approach is also discriminative point.

REFERENCES

[1] A. R. Metke and R. L. Ekl, "Security technology for smart grid networks," *IEEE Transactions on Smart Grid*, Vol.1, No.1, pp.99-107, 2010.

[2] S. Sridhar, A. Hahn and M. Govindarasu, "Cyber-physical sysytem security for the electric power grid," *Proceedings of the IEEE*, Vol.99, No.1, pp.1-15, 2012.

[3] Y. Fujita and T. Namerikawa, "Synchronization Condition of Power Networks by Using Non-Uniform Kuramoto Model," *International Symposium on Nonlinear Theory and its Applications*, 2011.

[4] A. Domínguez-García and S. Trenn, "Detection of impulsive effects in switched DAEs with applications to power electronics reliability analysis," *Proc. of the IEEE Conference on Decision and Control*, pp.5662-5667, 2010.

[5] F. Pasqualetti, F. Dörfler and F. Bullo, "Cyber-Physical Attacks in Power Networks: Models, Fundamental Limitations and Monitor Design," *Proceedings of the IEEE Conference on Decision and Control*, pp.2195-2201, 2011.

[6] T. Namerikawa and T. Kato, "Distriburted Load Frequency Control of Electrical Power Networks via Interative Gradient Method," *Proc. of the 50th IEEE Conf. on Decision and Control*, pp.7723-7726, 2011.

[7] K. Kosugi, S. Tokumoto and T. Namerikawa, "Fault-tolerant Sensor Network based on Fault Evaluation Matrix and Compensation for Intermittent Observation," *Proc. of the 51st IEEE Conf. on Decision and Control*, 2012.(to be published)

[8] A. Teixeira, H. Sandgerg and K. H. Johansson, "Networked Control Systems under Cyber Attacks with Applications to Power networks," *Proc. of the American Control Conference*, pp.3690-3696, 2010.

[9] D. J. Klein and W. Xiao, "Resistance distance," *Mathematical Chemistry*, Vol.12, No.1, pp.81-95, 1993.

[10] A. Pasdar and S. Mirzakuchaki, "A Solution to Remote Detecting of Illegal Electricity Usage Based on Smart Metering," *Proc. of the IEEE International Workshop on Soft Computing Applications*, pp.163-167, 2007.