Identification of Compromised Power System State Variables

Nathan Wallace, Stanislav Ponomarev, and Travis Atkison
Deptartments of Electrical Engineering, Cyber Engineering, and Computer Science
Louisiana Tech University
501 Dan Reneau Drive Ruston, LA 71272
Email: dcsl@latech.edu

***Abstract*—Securing the critical infrastructure power grid is one of the biggest challenges in securing cyberspace. In this environment, control devices are spread across large geographic distances and utilize several mediums for communication. Given the required network topology of the power grid several entry points may exist that can be utilized for compromising a control network. This article explores a cyber event detection scheme based on the Grubbs' test to classify univariate values. The test is conducted only after a power system instance has been classified as containing a cyber-event. The classification of each instance is made via principal component analysis and the Hotelings T^2 value. A Monte Carlo simulation is used to determine a set of converging power system instances and is based on the Newton-Rhapson method to solve the power flow equations of a 5 bus power system. Results indicate successful classification at a rate of 90%.**

***Index Terms*—SCADA, PLC, control systems, state estimation, intrusion detection**

I. Introduction

Perhaps one of the biggest challenges of securing cyberspace, is the ability to secure the critical infrastructure power grid. This is in part due to the inter-connective nature of the power grid and how every aspect of modern life is driven by the notion of always having power available. The power grid is composed of a meshed network of geographically distributed industrial control systems (ICS) that span large distances and utilize multiple communication mediums and protocols. Such interlacement, unbeknownst to the utility provider or independent system operator (ISO), can provide an individual or nation-state with malintent direct access to the control local area network. Once the control LAN has been breached, control decisions can be made that are outside the intended operation specifications, the most harsh being a full denial of service attack. The critical infrastructure power grid has recently seen an increase in the implementation of networked solid state devices. The key goal of such influxes is to increase the number of reporting nodes in the Wide Area Measurement System (WAMS) for the purpose of billing, state estimation, grid health, and for the efficient delivery of electricity to its consumers. However, security becomes a concern when the control decisions being implemented in the power system are based on the values being reported by the nodes in the WAMS.

In a recent effort known as Project Shine, over 7,200 control devices were found to be directly connected to the World Wide Web [1]. These startling results indicate that critical control devices have and will continue to be accidentally connected in a manner that is inconsistent with the so called 'air-gap' separation. Other possible and, in some cases, historically documented breaches into power systems are conducted via insider threat, the use of a zero-day attacks, or unpatched system attacks. The approach presented in this article aims at solving the detection of attacks against power systems using a context specific approach.

The approach presented in this article uses the Grubbs' Test to identify the reporting power system node that was compromised. This analysis is made possible by first using a transformation that identifies if an instance contains a cyber-event. Specifically, principal component analysis is used as the approach for transforming power system instances, and the Hotelling T^2 metric is used for the classification of each newly observed instance. Once an instance is labeled as suspect, the state parameters contained within that instance are compared against the variances of previously observed or trusted state parameters using the normalized residual test, Grubbs' Test, in an effort that identifies the node or control device that was the target of the intrusion. The identification scheme is applied to the data resulting from a Monte-Carlo simulation using an iterative solution to the power flow equations. The iterative solution used for the development of system data is the Newton-Rhapson method and is known to be the most common approach for solving the power flow equations [2].

Details and model assumptions of the power grid are described in Section II and Section III. An overview of the residual test, Grubbs' test, and the dimensional transformation technique, principal component analysis (PCA), is given in Section IV. The cyber-event model outlined in Section III-B describes how the instances are created such that they represent a possible malicious attack on the power system or a failed sensor. Lastly, the results of the cyber-event detection scheme are presented in Section V followed by future work and conclusions.

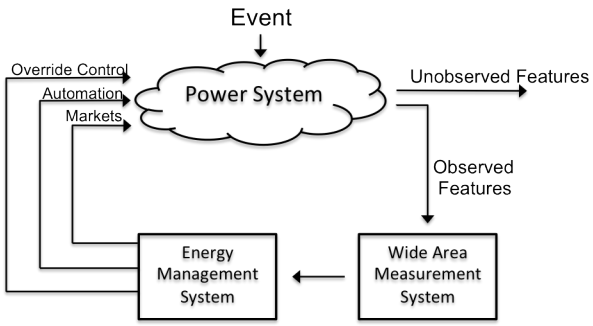

Fig. 1. Basic Power System Application Feedback Model

II. The Power Grid

A basic model, derived from similar ones presented in [3], [4], of a power system application with state feedback is presented in Figure 1. The feedback model shown is governed by the energy management system (EMS) which during an event will instruct the SCADA system to send control commands to the power system application [5]. An event can consist of a fault, i.e. a down power line, or a disturbance as modest as a customer turning on a lamp. Events change the operating conditions of the application and, if drastic enough, will cause the EMS to take immediate action to protect the system from catastrophic failure. In instances where the event does not cause immediate harm to the power system the EMS will remain idle or change control parameters to more economically provide power to customers. The purpose of this feedback interface is for constant monitoring and control of the power system application in an effort to ensure the constant and stable generation and delivery of power.

The primary steady state algorithms that determine the stability and reliability of the *critical infrastructure power grid* are: 1) Power Flow, 2) Optimal Power Flow, and 3) State Estimation. The power system challenge is to try to solve the nonlinear power balance equations in near real time given a percent of system values. The system state uses Kirchoff's Law at each power system bus throughout the system in question. Kirchoff's Law states that the sum of the powers entering a bus must be zero. The active and reactive components of the power flow equations in polar representation form from bus i to bus j can be determine by solving Equations 1 and 2.

$$0 = \Delta P_i = P_i^{injec} - V_i \sum_{i=1}^{n} V_j Y_{ij} cos(\theta_i - \theta_j - \varphi_{ij}) \quad (1)$$

$$0 = \Delta Q_i = Q_i^{injec} - V_i \sum_{i=1}^{n} V_j Y_{ij} sin(\theta_i - \theta_j - \varphi_{ij}) \quad (2)$$

where, P_i^{injec} and Q_i^{injec} are the injected powers into each bus, V_i is the voltage on bus i and Y_{ij} is element ij of the admittance matrix. Optimal power flow is the result of finding the desired power system state variables based on one or multiple cost functions. Examples of cost functions include minimization on power losses and fuel costs of generation. State estimation describes the process of estimating the state of the power system based on an incomplete picture of the system being observed. With state estimation, system parameters are measured using intelligent electronic devices (IEDs) and are reported back to the SCADA system.

A. State Estimation and Power Flow

Power flow analysis uses an iterative method, in most cases the Newton-Rhapshon method [2], for solving the nonlinear algebraic power flow equations, Equations 1 and 2 [7]. Convergence is said to happen when the error or mismatch drops below a certain threshold. For instance, the error stopping point used in this approach is $\varepsilon_s = 0.01$. This means that the absolute values of both the active and reactive power mismatches all had to be below 0.01 to be considered a converging instance. Also, for this examination convergence had to occur within 15 iterations or the instance was declared a non-converging instance. On average the 5 bus systems converged within 4 iterations. The extreme of 15 iterations was selected as a stopping point given that if the system did not converge within 15 iterations it is likely for that given set of inputs the system cannot exist. The fact of non-convergence corresponds to the likelihood that the power system being observed does not exist at that given set of inputs. For a more detailed description of the iterative solutions to the power flow problem the reader is encouraged to view the following referenced text [2], [7], [8].

III. Simulation Models

A. Power System Model

To demonstrate the identification of cyber-events a relatively simple power system was selected. Multiple instances of this model were conducted using the Newton-Rhapson method to solve the nonlinear algebraic power flow equations. Using the 5 Bus power system [8] shown in Figure 2 a series of power flow simulations were conducted. The system shown is a 100 MVA 138 kV system with the swing Bus positioned at Bus #1 or the Slack Bus. Generators are connected at Bus #1 and Bus #2. Loads are connected to every Bus in this model and are identified by that Bus's number. Table I shows the impedances used for the six transmission lines considered in this system model. A snapshot of the Bus input data is shown in Table II. This information serves as the input parameters to the power flow equations and with the successful convergence of the Newton-Rhapson method the other variables can be determined. Bus #3 is a voltage controlled Bus and is part of the input variable set. The slack Bus is simulated in such a way that given the inputs shown in Figure 2 it picks up the remaining slack to supply the required load.

B. Cyber-Event Model

The cyber-event model used for this detection approach is two-fold in that it represents two possibilities that can occur

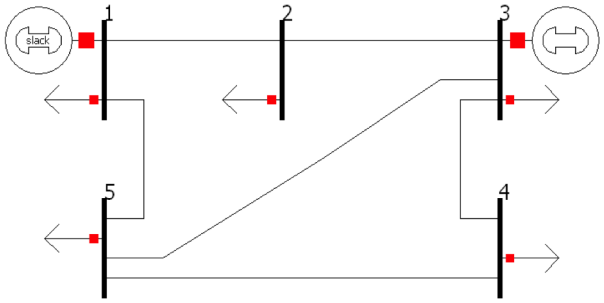

Fig. 2. Five Bus One Line Diagram [8]

TABLE I
5 BUS TRANSMISSION LINE PARAMETERS [8]

Bus - Bus	Line Length (mi)	R	X	B
1 - 2	40	0.042	0.168	0.041
2 - 5	30	0.031	0.126	0.031
2 - 3	30	0.031	0.126	0.031
3 - 4	80	0.084	0.336	0.082
3 - 5	50	0.053	0.210	0.051
4 - 5	60	0.063	0.252	0.061

in a power system. Event #1 can be considered to be a non-malicious incident in which the controller or sensor in the field making the measurement breaks or becomes damaged as a result of natural causes. Some examples of this may include natural disasters, faulty equipment, or wear on the device over the years. Event #2 can be classified as an actual malicious event in which an attacker purposely launches an attack against the control system. Examples of this include the falsification or spoofing of data values reported from a smart meter as revealed by Brinkhaus et al [9]. This work currently makes no distinction of the two events only that it is able to determine that an event occurred. Once detection has occurred that instance then can be further investigated and the actual cause of the event can be determined.

The approach presented in this article assumes that both Event#1 and Event#2 will produce a state value of zero at the origin of the event. This assumption provides an initial starting point for the development of the detection scheme presented in this article. Furthermore the cyber-event model assumes that only one cyber-event occurs per instance and hence forth makes no distinction between the two events based on the developed identification scheme. An alarmed instance will only show that either event could have been the cause of the cyber-event.

To simulate these types of events a random instance from data matrix $\mathbf{X}$ was selected. This random instance vector $\overrightarrow{X_r}$ serves as the basis for the event simulation. Currently a total of ten events are simulated each event corresponds to an instance or row in a new suspicious data set $\mathbf{X}'$. For the first event, first row in the suspicious set, the variable x_1 of $\overrightarrow{X_r}$ is changed to a zero representing either a failure or an attack occurring at the voltage reading on Bus #1. This is done while holding all other values equal to the corresponding variables of the random vector $\overrightarrow{X_r}$. For each subsequent event instance the next variable is changed while holding all variables consistent with the values from $\overrightarrow{X_r}$. All simulated events occur at the bus voltages and the real power at each of the 5 loads.

TABLE II
5 BUS INPUT SNAPSHOT

Bus #	Type	V	Delta	PG	QG	PL	QL
1	0	-	0	-	-	0.65	0.3
2	1	-	-	0	0	1.150	0.6
3	2	1.020	0	1.8	-	0.7	0.4
4	1	-	-	0	0	0.7	0.3
5	1	-	-	0	0	0.850	0.4

IV. IDENTIFICATION APPROACH

A. Principal Component Analysis

In any determinable system there is a finite number of driving forces which governs how the system behaves. By observing grouping phenomenon in the data it is possible to replace a group of variables with a single new variable, greatly reducing the redundancy in the data. Principal component analysis (PCA) is a quantitative process for achieving a system simplification. A decrease in redundancy and an overall simplification of the data is made possible through a transformation into a new vector space where all the basis vectors are independent of each other. The basis vectors in the new dimensional space are called principal components [10]. PCA is based on the statistics of a training set to linearly transform the set in such a way that the new primary basis are independent of each other. The linear transformation used is based on a covariance matrix which is defined by the patterns found in the training set. PCA finds a linear transformation such that

$$\mathbf{Y} = \mathbf{W}\mathbf{X} \tag{3}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are mxn matrices related by a transformation $\mathbf{W}$. Based on Equation 3 the following variables can be defined: $\mathbf{w_i}$ are the rows of $\mathbf{W}$, $\mathbf{x_i}$ are the columns of $\mathbf{X}$, and $\mathbf{y_i}$ are the columns of $\mathbf{Y}$.

The row vectors of $\mathbf{W}$ $\{w_1, ..., w_m\}$ are called the principal components of $\mathbf{x}$. Before PCA can be applied to a data set it is customary to first preform sanitization on the data. This sanitization guarantees any unintended biassing of the new components. After centering the normalized covariance $\mathbf{S_X}$ was determined using the unbiased estimator for normalization.

$$\mathbf{S_X} = \frac{1}{n-1}\mathbf{X}\mathbf{X}^T \tag{4}$$

This produced a covariance matrix with dimensions mxm with the diagonal terms representing the variances and off-diagonal terms representing the covariances of data matrix $\mathbf{X}$. The closer the off-diagonal terms are to zero the closer the variables represented by the indices of $\mathbf{S_X}$ are to being completely uncorrelated. Conversely, the higher these off-diagonal terms are the more correlated the two variables are. Also the higher the off diagonal terms are the higher the redundancy is in the data matrix $\mathbf{X}$.

The linear transformation produced by PCA selects a transformation $\mathbf{W}$ such that the principal components or basis vectors w_i produced are completely orthonormal. Orthonormality is ensured due to the fact that the dot product of each basis vector with another produces the Kronker delta function, $w_i \cdot w_j = \delta_{ij}$. In addition to being orthonormal, the basis vectors are ordered based on the amount of variance that is being accounted for by that basis vector or principal component. This corresponds to the fact that PCA will produce a transformation matrix $\mathbf{W}$ such that the variance of data matrix $\mathbf{X}$ is mostly accounted for by principal component w_1. As hinted at in the previous section the lower the diagonal terms of the covariance matrix are the lower the redundancy is in the data. Therefore the solution to PCA seeks a covariance matrix $\mathbf{S_Y}$ such that the off-diagonal terms are zero where,

$$\mathbf{S_Y} = \frac{1}{n-1}\mathbf{Y}\mathbf{Y}^T \tag{5}$$

Plugging Equation 3 into Equation 5 we have

$$\mathbf{S_Y} = \frac{1}{n-1}\mathbf{W}(\mathbf{X}\mathbf{X}^T)\mathbf{W}^T \tag{6}$$

With this solution to PCA it can be shown that the principal components of data matrix $\mathbf{X}$ are the eigenvectors of $\mathbf{X}\mathbf{X}^T$ or are the rows of $\mathbf{W}$. Also, the i^{th} diagonal term of $\mathbf{S_Y}$ is the variance of $\mathbf{X}$ projected onto $\mathbf{p_i}$.

B. Classification of New Power System Instances

The Naive Bayes classifier Hotelling T^2 metric, $T^2 = n(\mathbf{X}-\mu)'\mathbf{S}^{-1}(\mathbf{X}-\mu)$, is utilized for detection and is an extension of the t-test used to determine the difference between means of two independent variables. This extension allows for a statistical measure of the multivarite distance of each instance from the center of the data set in the reduced dimensional space. The result allows for the detection of instances that occur at far distances from the data center as defined by data matrix $\mathbf{X}$. The detection approach presented in this article is a probabilistic approach in describing how likely an instance is to occur. Instances that fit to the dynamics of the data matrix $\mathbf{X}$ or control set have a high likelihood of occurring while instances that lie on the boundaries are less likely to occur. It can also be shown that the Hotelling's T^2 value follows the $\mathcal{F}$ distribution as defined by Equation 7 [11]

$$T^2 \sim \frac{(n-1)p}{(n-p)}\mathcal{F}_{p,n-p}(x) \tag{7}$$

where p is the number of principal components retained and n is the number of instances in the sample space. Because over 90% of the variance is accounted for by the first 8 principal components, a value of $p = 8$ was used. The $\mathcal{F}$ cumulative probability distribution function returns the cumulative probability of obtaining a value x for given parameters p and n. Rearranging Equation 7 we can calculate that the probability of observing at least T^2 is $P(\geq T^2) = 1 - F_{p,n-p}(z)$ where,

$$z = T^2\frac{(n-p)}{p(n-1)}$$

This allows for a probabilistic metric to determine whether or not an instance is in control. If the instance is in control then it follows the dynamics as defined by the data matrix $\mathbf{X}$. Using the maximum Hotelling T^2 value as a threshold all newly observed power system instances are classified as either suspect or non suspect. The smaller the value the closer the power system instance aligns with the dynamics of the trusted model. Then upon classification a control engineer can perform further analysis to determine the root cause of the cyber-event.

C. Grubbs' Test

The Grubbs' test, also known as the maximum normed residual test, is used to detect outliers in a univariate data set [12] [13]. Formally the test can be defined as a means of hypothesis testing. Using the test statistic G as defined by Equation 8 the result of the hypothesis test can either be H_0 for *no outliers in the data set* and H_a if there is exactly one outlier in the data set.

$$\mathbf{G} = \frac{max|Y_i - \bar{Y}|}{s} \tag{8}$$

With Y_i representing the measured value, $\bar{Y}$ representing the sample mean, and s representing the standard deviation of the state variable it is possible to also define the critical region for each variable. The test hypothesis H_a is true if for a given data set $Y = [y_1, y_2, \ldots, y_{N-1}, y_N]$ Equation 9 holds true; with $t_{\alpha/(2N),N-2}$ denoting the critical value of the t distribution with (N-2) degrees of freedom and a significance level of $\alpha/2N$.

$$G > \frac{(N-1)}{\sqrt{N}}\sqrt{\frac{(t_{\alpha/(2N),N-2})^2}{N-2+(t_{\alpha/(2N),N-2})^2}} \tag{9}$$

For clarity the Grubbs' test is syntactically adjusted to fit the application of detecting the compromised power system state variable. For an incoming power system instance X_i, the dimensional transformation scheme, PCA, transforms it into a new vector space and a distance classifier is used to determine the validity of the instance. However, this does not identify the variable that was the source of the cyber-event. Therefore, after each power system instance cyber-event classification the Grubbs' test can be performed on each variable independently to determine any potential anomalies in that state variable based on historical readings. Each newly observed instance i is comprised of n variables with each variable labeled as $x_{i,j}$. By letting $Y_j = [x_{1,j}, x_{2,j}, \ldots, x_{N-1,j}, x_{N,j}]$ a new notation can be defined for the identification scheme. For instance the vector Y_1 describes the full set of bus 1 voltages.

Since the newest power system instance, if classified as containing a cyber-event, is the one under inspection with the Grubbs' test it is the N^{th} observation that will be calculated and compared. The Grubbs' test can now formally be defined as Equation 10 and 11. In Equation 11 a discriminate $\delta_{\alpha,N}$ is created that equals the right hand side of Equation 9.

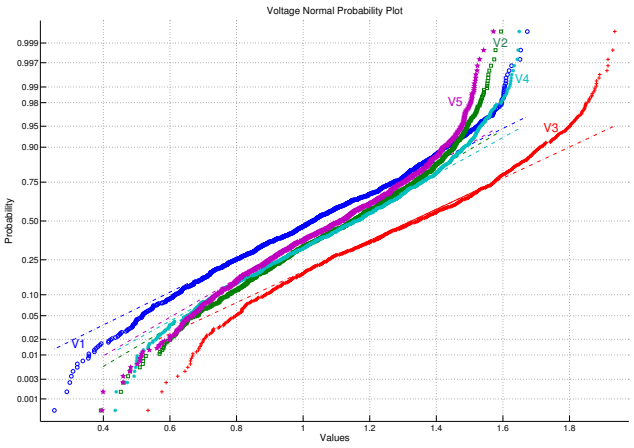

Fig. 3. Bus Voltages Distribution

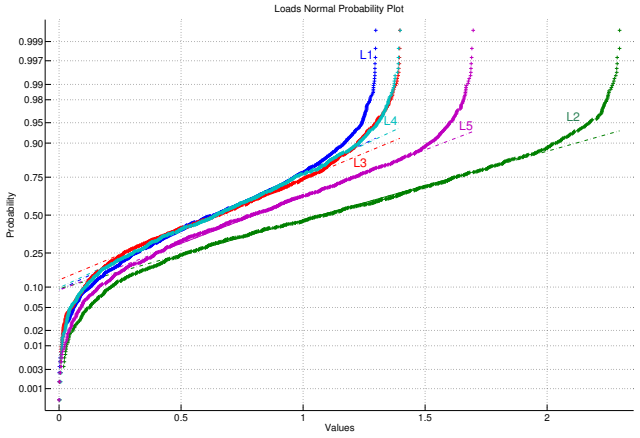

Fig. 4. Bus Loads Distribution

$$\mathbf{G} = \frac{|Y_{N,j} - \bar{Y}_j|}{s(Y_j)} \tag{10}$$

$$\mathbf{G} > \delta_{\alpha,N} \tag{11}$$

V. Event Classification

If a cyber event has occurred it is desired to detect such an event and be able to alert on intrusion or failure. This classification capability includes the identification of the compromised node. The immediate feedback will allow the trigger of an alarm allowing a security analyst or control engineer to further investigate the event. To better understand the power system state parameters trying to be secured Figure 3 and Figure 4 show the normal probability plots for the bus voltages and bus loads respectfully for a total of 996 converging power system instances.

Given that we now have defined a transformation matrix $\mathbf{W}$ such that this transformation has eliminated all redundancy when mapped to the dimensional space we can now interpret new instances of the power system. With a trusted model derived from known instances a threshold value, T^2_{thr} was utilized to classify newly observed power system instances and is based on the maximum Hotelling T^2 of the trusted model in the transformed dimensional space. Using a trusted model containing 996 simulated power system instances, the maximum threshold value was determined to be $T^2_{thr} = 332$. When each of the 10 simulated cyber-events were mapped to the new dimensional space as a single score, that event's Hotelling T^2 value was calculated. The T^2 value calculated for each power system instance containing a cyber-event is shown in Table III. This table reveals that the each power system instance that contained a cyber-event was successfully classified as such. However, the challenge then comes to classify the node or source of the cyber-event.

Using Grubbs' test, Equations 10 and 11, a classification was conducted within each power system state variable based on the outlier hypothesis testing. Recall that the 10 simulated cyber-events correspond to malforming trusted instances by changing only one of the variables to zero at a time. For a N value of 997 the discernment value δ for $\alpha = 0.5$ is 3.4705, $\delta_{0.5,997} = 3.4705$. Similarly the identification scheme was determined using a $\alpha = 0.5$ for N = 997, resulting in a discernment value δ of 4.039, $\delta_{0.05,997} = 4.039$. Using the calculated discriminant values combined with Equation 11 for classification, each variable can be identified as being the source of the cyber-event. The results of the identification for each discriminant value across all 10 power system state variables is shown in Table III. A 'Y' denotes that the associated discriminant function $\delta_{\alpha,N}$ successfully identified the source of the cyber-event, and a 'N' denotes a non-successful identification.

TABLE III
Identification Results N=997

j	Description	T^2	$\mathbf{G}$	$\delta_{\alpha=0.05,N}$	$\delta_{\alpha=0.5,N}$
1	Bus Volt	852.31	6.266	Y	Y
2	Bus Volt	960.50	8.005	Y	Y
3	Bus Volt	909.86	7.148	Y	Y
4	Bus Volt	954.85	7.364	Y	Y
5	Bus Volt	976.88	7.850	Y	Y
6	Load 1	994.10	4.089	Y	Y
7	Load 2	995.92	3.056	N	N
8	Load 3	995.62	3.769	N	Y
9	Load 4	995.81	3.869	N	Y
10	Load 5	995.99	3.545	N	Y

Results indicate that for both α values, $\alpha = 0.5$ and $\alpha = 0.05$, every simulated cyber-event on the bus voltages were identified. This perhaps could have been anticipated by thoroughly analyzing Figure 3 where it is observed that a majority of the previously observed power system bus voltages occur within the region $0.6 < V_i < 1.5$. This however is not the case for the bus loads. The real power of the bus loads seem to concentrate between the region $0 < L_i < 1.2$, with a steep descent towards 0. Therefore, a cyber-event of zero on a bus load will be harder to detect than a cyber-event of zero on the bus voltages. By changing the significance value α, a larger region is

covered inevitably increasing the classification potential of the discriminant classifier. However, this may lead to higher false positive identification. Using a higher alpha value, specifically $\alpha = 0.5$, all but one of the simulated cyber-events were identified.

VI. Future Work

Though the simulated cyber-events offer insight into a possible detection and identification scheme based on the Grubbs' test a more complete analysis of this approach would include a full mapping of detectable regions for cyber-events of varying values. One benefit of the extensive analysis would include the fact that regions of stealthiness can be mapped out for each variable. Furthermore, future work includes a weighted alpha value that changes depending on the variance found within each power system state variable. Such a technique may decrease the false positive rate of the detection and identification scheme.

VII. Conclusion

Using a normalized residual test, power system state variables were successfully identified as being the source of a cyber-event. The residual testing scheme utilized is a slight modification of the Grubbs' test to classify the newly observed power system state variables. Cyber-events are simulated by changing each power system bus voltage and the real power consumed at each bus independently to zero. A change of zero may be the result of a faulty equipment or an individual spoofing power system variables in an effort to lower his utility bill. The new observation was successfully classified as containing a cyber-event using a dimensional transformation to transform observed power system instances a probabilistic metric. Once the instance is found to contain a cyber-event the Grubbs' test was conducted to determine the power system state variable that was the source of the event. Such an analysis will allow the security investigator or control engineer to immediately isolate and fix the intrusion or problem.

The identification scheme described in this article is performed on a 5 bus power system. Trusted instances of the power system were determined using the Newton-Rhapson method of mismatch error less than 0.01 and convergence was required within 15 iterations. Principal component analysis (PCA) was used as a feature reduction method transforming 47 power system state variables into 8 principal components. Classification of each power system instance was based on a threshold Hotelling's T^2 value and if determined to contain a cyber-event the modified Grubbs' test was performed. This approach successfully classified 100% of the simulated cyber-event instances as containing a cyber-event and was able to identify 90% of the compromised power system state variables.

Acknowledgment

This research was supported by a Louisiana Board of Regents Graduate Fellowship.

References

[1] ICS-CERT, "Project shine," *ICS-CERT Newsletter Monthly Monitor*, vol. October-December, 2012.
[2] *Power Systems (The Electric Power Engineering Hbk, Second Edition)*. CRC Press, 2007.
[3] Z. Lukszo, *Securing electricity supply in the cyber age : exploring the risks of information and communication technology in tomorrow's electricity infrastructure*. Dordrecht New York: Springer, 2010.
[4] L. Grigsby, *Power system stability and control*. Boca Raton, FL: CRC Press, 2007.
[5] T. Gonen, *Electric power distribution system engineering*. New York: McGraw-Hill, 1986.
[6] G. Anders, *Probability concepts in electric power systems*. New York: Wiley, 1990.
[7] J. D. Glover, M. S. Sarma, and T. Overbye, *Power System Analysis and Design, Fifth Edition*. Cengage Learning, 2011.
[8] W. D. Stevenson, *Elements of Power System Analysis (Mcgraw Hill Series in Electrical and Computer Engineering)*. Mcgraw-Hill College, 1982.
[9] C. D. Brinkhaus S., "Smart hacking for privacy," 2011.
[10] K. J. Cios, W. Pedrycz, R. W. Swiniarski, and L. A. Kurgan, *Data Mining: A Knowledge Discovery Approach*. Springer, 2007.
[11] W. K. Hardle and L. Simar, *Applied Multivariate Statistical Analysis*. Springer, 2012.
[12] F. Grubbs, "Procedures for detecting outlying observations in samples," *Technometrics*, vol. 11(1), pp. 1–21, 1969.
[13] W. Stefansky, "Rejecting outliers in factorial designs," *Technometrics*, vol. 14, pp. 469–479, 1972.