

# Improving Attack Detection in Power System Control Center Critical Infrastructures Using Rough Classification Algorithm

Coutinho, M. P., Lambert-Torres, G., *Member*, IEEE, da Silva, L. E. B., *Member*, IEEE, Lazarek, H.

**Abstract**—Nowadays, Power System Control Centers are moving toward distributed and decentralized operations. The use of information technology (IT) to achieve service these goals produces vulnerabilities and security threats. To safeguard against the threat of cyber-attacks, service providers also need to maintain the accuracy, assurance and integrity of their interdependent data networks. This paper presents a novel technique for improving the security of Power System Control Centers in the Electric Power System Critical Infrastructure by implementing anomaly detection methods to identify attacks and faults. By using Rough Sets Classification Algorithm, a set of rules can be defined to the anomaly detection process. This can be used for identify attacks and failures and, also, for improving state estimation.

**Index Terms**—Critical infrastructure protection, Electric Power System, Power System Control Center, SCADA , Detecting Attacks, Rough Set Theory, Data Mining.

## I. INTRODUCTION

THE Control Center is the central nerve system of the Electric Power System Critical Infrastructure. The restructuring of the power industry has transformed its operation from centralized model to coordinated decentralized decision-making model [1]. In order to achieve the new model, the electricity management network, corporate network and the use of information technology (IT) produces vulnerabilities and exposes the electricity cyber infrastructure to securities threats [2,3,4,5,6]. In [7] are listed some

current initiatives for Security of the Critical Infrastructures.

As part of the Power System Control Centers, SCADA systems and Energy Management Systems (EMS) play a vital role in order to monitoring the safety, reliability, and protective functions of the power grid. However, these systems, that were designed to maximize functionality with little attention paid to security, represent potential vulnerability to disruption of service or manipulation of operational data that could result in public safety concerns [8].

According to [9] there are two approaches to become SCADA systems more secure: One is to identify problems at the perimeter of the system using anti-virus and Intrusion Detection Systems (IDS). The second is to model the normal data flows and control operations within the SCADA system to detect anomalies caused by attempts to change or damage the system.

Using the second approach, this paper presents the development of the technique for implementing anomaly detection to monitor Power System Control Centers, previously introduced in [7], where the problem was addressed using Rough Sets Classification Algorithm proposed by Pawlack et al [10].

This paper is organized as follow: Firstly, an overview of the Electric Power Systems Critical Infrastructure, SCADA systems and Rough Sets Classification Algorithm is presented. Then, the architecture of the Anomaly Detection System is introduced and the methodology to build the knowledge data base and how to extract the rules from such data base is described. A Six Bus Power System is used as an example.

---

This work was supported in part by the Brazilian Research Council (CNPq) and Minas Gerais State Research Foundation (FAPEMIG).

M. P. Coutinho, G. Lambert-Torres, and L.E. Borges da Silva are with the Federal University of Itajuba (UNIFEI), Itajuba, MG, 37500-903, Brazil (phone: +55-35-36291240; fax: +55-35-3629118755; e-mail: {maurilio.coutinho, germanoltorres}@gmail.com).

H. Lazarek is with the Technische Universität Dresden, Dresden, Germany.

## II. ELECTRIC POWER SYSTEM CRITICAL INFRASTRUCTURE

In general, an Electric Power System Critical Infrastructure is highly interconnected and dynamic, consisting of several utilities. Due to its hierarchical organization, it is sub-divided into regional grids. Each sector is further split into generation, transmission, distribution, and customer service systems, supplemented with an energy trading system. The Power Grid is comprised of a myriad assets, such as Generation Plants, Transmission Lines, Transmission and Distribution Power Substations, Local, Regional and National Power System Control Centers, Remote Terminal Units (RTUs)/Intelligent Electronic Devices (IEDs), and Communication Links [11].

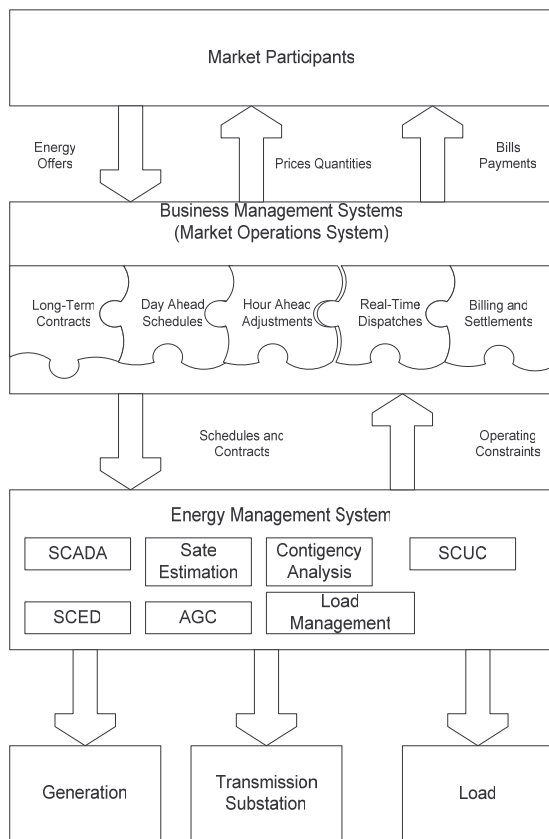


Fig. 1. Power System Control Center Interactions [1]

Accordingly [1], the Power System Control Centers are characterized by: Electric Management Systems, which allow operators to regulate power flow, the Supervisory Control and Data Acquisition (SCADA) systems for monitoring the safety, reliability, and protective functions of the power grid, and the Business Management Systems (BMS)

responsible for business application (market operations). See Fig. 1.

## III. PROTECTING SCADA SYSTEMS

Accordingly [1], a Power System Control Center has several networks: a master-slave network from Remote Terminal Units (RTUs) to the control center with dedicated physical links, a LAN from EMS application servers, a point-to-point network for inter-control center connections, and the Internet for BMS market functions. See fig. 2.

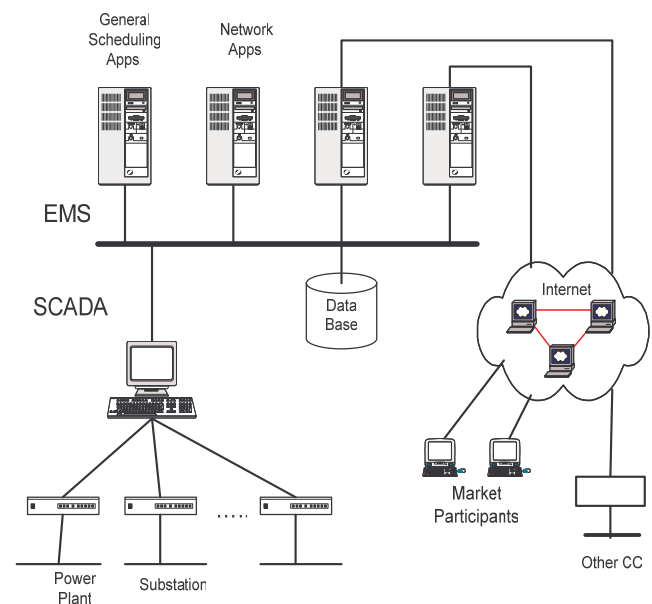


Fig. 2: Control Center Architecture [1]

The SCADA system are used for data collection from sensors and instruments located at remote sites, like the substation control center and the power plant, to transmit and to display the data at a Power System Control Center for Control and Supervisory purposes. The SCADA system can monitor and control hundreds of I/O points. The RTUs are located between the remote sensors and the Control Center in order to gather the data from sensors and field devices. The sensors have Digital or Analog I/O and these signals are not in a form that can be easily communicated over long distances. Therefore, the RTUs digitize the sensor signals so that they can be digitally transmitted via communication protocols over long distances to the Control Center.

The SCADA Communications can employ a

diverse range of both wired (leased lines, dialup line, fiber, ADSL, cable) and wireless media (spread spectrum, cellular, WLAN or satellite). The choice depends on a number of factors that characterizes the utility existing communication infrastructure.

In order to analyse the vulnerabilities of a SCADA system communication model, as showed in the figure 3, it is pointed out various weak points where insider or outsider attackers can get access to the SCADA Master and the RTU. For example the Circuit Breaker can be considered an attack object because of the Internet connectivity via corporate network or via remote access using public telephone network. In the case of a success access (inside or outside) to the RTU, two possible scenarios can be visualized: (1) The attacker assumes the control of the circuit breaker or (2) the attacker corrupts the information collected by the RTU. In order to detect such scenarios, anomaly detection techniques are used to identify these threats as well as the type of attack.

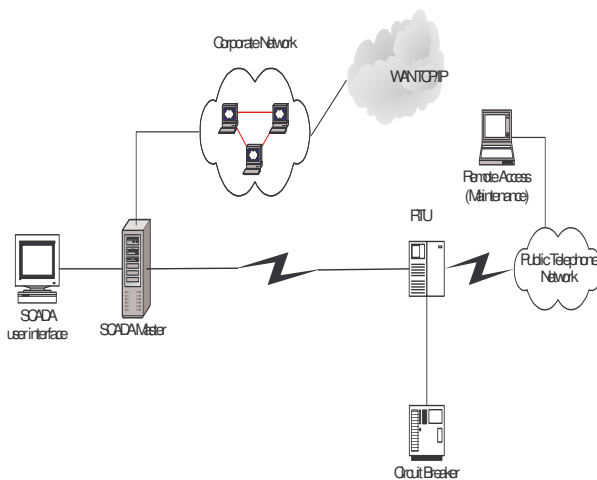


Fig. 3: SCADA System Communication Model

The Intrusion Detection System (IDS) has been studied widely in recent years. Anomaly-based IDS discovers attacks by identifying unusual behaviour (anomalies) on a host, network or application. They function on the observation that some attackers behave differently than “normal” users or events and thus, can be detected by systems that identify these differences. An extended bibliography on IDS is presented in [7].

#### IV. PROBLEM DEFINITION

The operation of a power system is intrinsically complex due to the high degree of uncertainty and the large number of variables involved. The various supervision and control actions require the presence of an operator, who must respond efficiently to the most diverse requests, by handling various types of data and alarm information.

As seen before, the data and information come from measurements of SCADA system or from the computational processes. The size of the current database in a Power System Control Center has increased tremendously over the past few years due to the use of network communications, which renders their control systems more vulnerable to manipulation by malicious intruders. In order to improve the security of the Power System Control Centers, anomaly detection can be used to identify corrupted values caused by malicious attacks and faults.

The system operator must be apprised of the current state of the system and some forecasted position, such as load forecasting, maintenance scheduling, in order to take a control action (switching, changing taps, and voltage levels). One of the most important operator tasks is to determine the current operational state of the system. To accomplish this task, the operator receives many data from/into the system. By handling these data, the operator attempts to build an image of the operation point.

The analysis attempts to assess the operational mode in one of the 2 states: normal, or abnormal. In the first state, normal, all loads are supplied and all measurements are inside the nominal rates. In the second state, abnormal, all loads continue to be supplied, but some of the measurements are outside the nominal rates or some loads are not supplied. The operator must regularly analyze the system security, even when the operation state is normal. This analysis is conducted according to possible contingencies that could affect the power system.

According to Xuan Jin et al [12], there are a number of ways in which anomaly-detecting methodologies could enhance the integrity and security of electricity data. Firstly, it could act as a

useful complement to existing techniques, such as state estimation, for verifying the likely correctness of electricity measurements and give operators constant feedback about changes the integrity and reliability of the data. A second application is the improvement of standard protection devices such traditional IDS and virus checker.

## V. ROUGH SETS CLASSIFICATION ALGORITHM

The Rough Set Theory, developed by Pawlak [10], has emerged as a mathematical method to manage uncertainties from inexact, noisy and incomplete information and it has been one of the focal point research areas in artificial intelligence since its advent [13]. In [14] it is presented the basic concepts of rough set theory and point out some rough set-based research directions and applications.

Before presenting the algorithm, two major concepts in the Rough Set theory, *reduct* and *core*, must be defined. These concepts are important in the knowledge of base reduction.

Let  $\mathbf{R}$  be a family of equivalence relations. The reduct of  $\mathbf{R}$ ,  $RED(\mathbf{R})$ , is defined as a reduced set of relations that conserve the same inductive classification of set  $\mathbf{R}$ . The core of  $\mathbf{R}$ ,  $CORE(\mathbf{R})$ , is the set of relations that appear in all reduct of  $\mathbf{R}$  (i.e., the set of all indispensable relations to characterize the relation  $\mathbf{R}$ ). The main idea behind the knowledge base reduction is a simplification of a set of examples. This can be obtained by the following procedure:

- a) Calculate the core of the problem;
- b) Eliminate or substitute a variable by another one; and
- c) Redefine the problem using new basic categories.

The algorithm that provides the reduction of conditions can be represented by the following steps:

**Step 1:** Eliminate dispensable attributes

**Step 2:** Compute the core of the set of examples.

**Step 3:** Compute the reduced set of relations that conserve the same inductive classification of the original set of examples.

**Step 4:** Merge possible examples and compose the final set of rules.

## VI. ANOMALY DETECTION ARCHITECTURE

The proposed solution for the attack scenarios pointed out in the Sections III and IV using anomaly detection is presented on figure 4 and uses intelligent techniques to extract knowledge from the SCADA system. The approach is divided in 2 steps: Firstly, the knowledge extractor should generate a set of rules that will determine the normal or abnormal behaviour of the system. The data come from RTUs and will be checked against the set of rules to define the normality of the measurements. Secondly, the anomaly detector should recognize the type of attack occurred.

In order to satisfy the limited SCADA Master computational resources, the proposed model should reduce the number of input variables and the number of examples, offering a more compact set of rules for the anomaly detector.

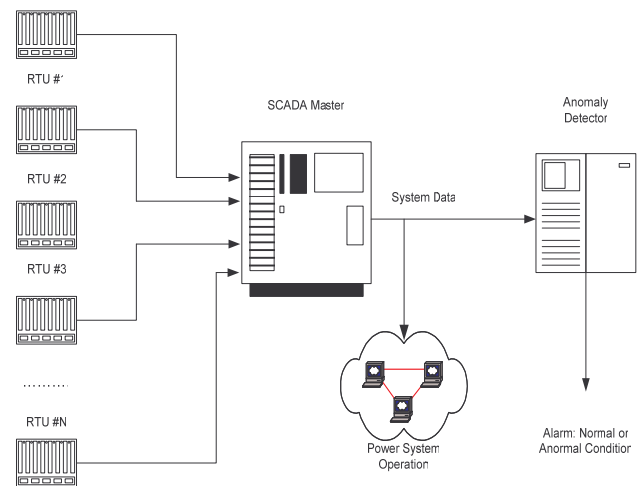


Fig. 4: Proposed Anomaly Detection Architecture

As previously defined in [7] the model uses the Rough Sets Theory to implement the data reduction. This technique is used mainly because:

- It reduces the number of rules without reducing the system knowledge base;
- It has a dynamic behaviour, because not informed rules by the expertise technician can be extracted from the system;
- It reduces the computational resources needed, like memory capacity and processor power;

However, the technique needs a huge amount of collected data to build the knowledge data base.

VII.EXPERIMENTS AND RESULTS

The diagram in the figure 5 represents the test environment for the proposed architecture. The main blocks depicted in the Fig. 5 are:

- Power Flow Module: To solve the power flow on an electric power system. This program was adapted from [15].
- SCADA Simulator Module: This program simulates the functions performed by an online Supervisory Control and Data Acquisition system. The idea is to simulate the power system network, calculates all the voltages, power flows and injections on the system and then associates these quantities with a specification of where the measurements are being made on the system. This program was adapted from [15].
- State Estimator Module: Program for state estimation process adapted from [15].
- Rough Set Rule Extractor Module: This Module extracts rules from the knowledge data base, using the Rough Sets Classification Algorithm
- Anomaly Detection System Module: This Module uses the rules defined by the Rough Sets Rule Extractor to determine the state of the SCADA Output data.

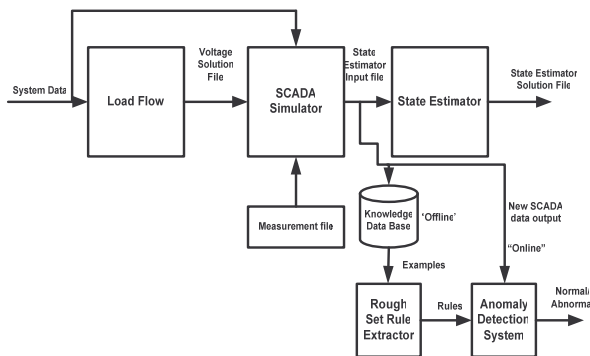


Fig. 5 Test Environment Diagram.

According [15], state estimation is the process of assigning a value to an unknown system state variable based on measurements from that system according to some criteria. According [16], “numerical estimation algorithms rebuild the state of the power system in case of missing and/or corrupted data: however this approach does not

address the problem of giving a normal/abnormal state assessment, and in some cases could tend to hide traces of an ongoing attack or of other anomalies”. This is a risky assumption since there are often configuration errors and there is always the chance that an attacker could be mediating between the control centre and the electricity network [9].

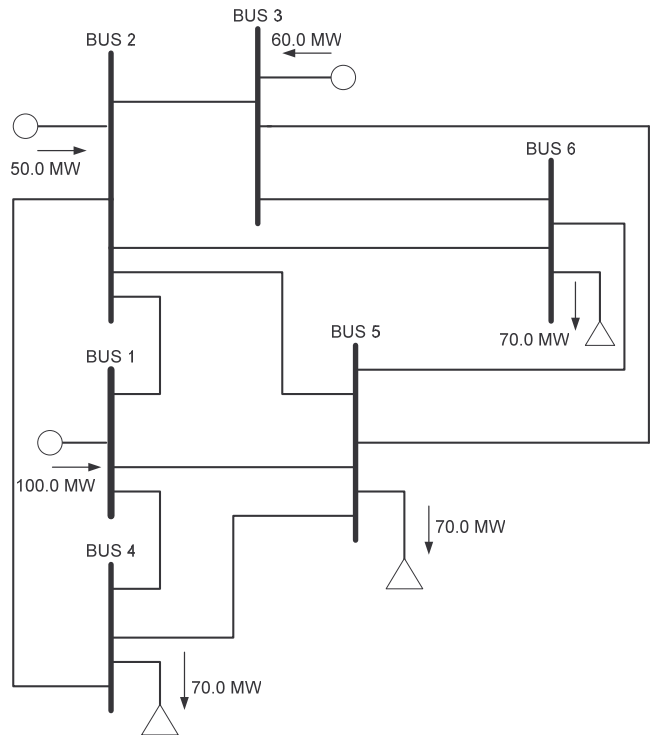


Fig. 6 Six Bus Power System [15].

To test the proposed anomaly detection model, it was used the environment test specified in the fig. 5 and a six bus network described in [15] and presented in the fig. 6. Test data was generated by introducing errors into the normal state estimation input file, generated by the SCADA Simulator Program. Xuan Jin et al [12] considered 5 types of corruption applied to the electricity data: constant bias with normally distribution deviation, loss of decimal point, sign switch, fixed at fixed value and fixed at random value. They attribute these errors to the fact that electricity measurements can be altered by random noise, attacks, software bugs, meter failures, EMI and transmission errors.

The ability of the proposed anomaly detection model to identify normal and abnormal conditions were evaluated against the ability of the state estimator module to provide a reasonable output.



while the state estimation process presented a possible output that could guide the operator to take wrong actions.

### VIII. CONCLUSIONS AND FUTURE WORK

Critical Infrastructures, such Electric Power Systems, are vital for our modern society. Therefore they require protection from a variety of threats, and their Power System Control Center Networks are potentially vulnerable to cyber attacks. The Anomaly Detection System is an important tool to increase the security of such Power System Control Center Networks. This paper presents an Anomaly Detection Model using a reduced set of rules extracted from a Electric Data Base Knowledge using Rough Set Theory. A test environment was proposed and implemented and an example for power system control centers demonstrated that the technique proposed has many advantages, such as simplicity of implementation and favorable performance. Future work includes expanding the error types introduced in the SCADA output file and the error type identification process. Besides it is the intention to compare such technique using the "Test Data for Anomaly Detection in Electricity Infrastructure" proposed in [17].

### REFERENCES

- [1] Wu, F. F., Moslehi, K., Bose, A., "Power System Control Centers: Past, Present, and Future", Proceedings of the IEEE, Vol.93, No.11, November 2005, pps. 1890-1908
- [2] Naedele, M., "IT Security for Automation Systems – Motivations and Mechanisms", ATP International, Vol. 1(1), 11/2003, <http://www.tik.ee.ethz.ch/~naedele/publications.html>
- [3] Schainker, R., Douglas, J., Kropp, T., "Electric Utility Responses to Grid Security Issues", IEEE Power & Energy Magazine, March/April 2006.
- [4] Geer, D., "Security of Critical Control Systems Sparks Concern", Computer, Vol. 39, Issue 1, January, 2006, pps 20-23.
- [5] Tani, M., "DOE Focuses on Cyber Security", Transmission & Distribution World, Vol 59, No. 3, March 2007, pps. 26-32.
- [6] Naedele, M., "Addressing IT Security for Critical Control Systems", 40th Hawaii Int. Conf. on System Sciences (HICSS-40) Hawaii, January 2007.
- [7] Coutinho, M.P., Lambert-Torres, G., da Silva L.E.B., Lazarek, H., "Detecting Attacks in Power System Critical Infrastructure Using Rough Classification Algorithm", Proceedings of the First International Conference on Forensic Computer Science, No.1, Vol.1, November 2006, pps. 93-99, Brasil.
- [8] Amanullah, M.T.O, Kalam, A., Zayegh, A., "Network Vulnerabilities in SCADA and EMS", 2005 IEEE/PES Transmission and Distribution Conference & Exhibition: Asia and Pacific, Dalian China.
- [9] Bigham, J., Gamez, D., and Ning Lu, "Safeguarding SCADA Systems with Anomaly Detection", V.Gorodetsky et al.(Eds.):MMM-ACNS 2003, LNCS 2776, pp. 171-182, Springer-Verlag Berlin Heidelberg, 2003.
- [10] Pawlak, Z., "Rough Sets", International Journal of Information and Computer Sciences, Vol.11, pp. 341-356, 1982.
- [11] Goetz, E., "Cyber Security of the Electric Power Industry", Institute for Security Technology Studies at Dartmouth College", December, 2002 .
- [12] Xuan Jin, Bigham, J., Rodaway, J., Gamez, D., Phillips, C., "Anomaly Detection in Electricity Cyber Infrastructure", Proceedings of CNIP, 2006, <http://www.davidgamez.eu/pages/publications.html>
- [13] Chengdong Wu, Yong Yue, Mengxin Li, Asei Adjei, "The rough set theory and applications", Engineering Computations, Vol. 21, No.5, 2004, pp 488-511, Emerald Group Pub. Limited, UK.
- [14] Pawlak, Z., Skowron, A., "Rudiments of Rough Sets", ScienceDirect, Information Sciences 177(2007)3-27, [www.sciencedirect.com](http://www.sciencedirect.com).
- [15] Wood, A.J., Wollenberg, B.F., "Power Generation Operation and Control", 2<sup>nd</sup> Edition, John Wiley & Sons, Inc., 1996.
- [16] Martinelli, M., Tronci, E., Dipoppa, G., Balducci, C., "Electric Power System Anomaly Detection Using Neural Networks", M.Gh. Negoita et al. (Eds.), KES 2004, LNAI 3213, pp. 1242-1248, 2004, Springer Verlag Berlin Heidelberg.
- [17] Bigham, J., Gamez, D., Xuan Jin, Chris Phillips, "Test Data for Anomaly Detection in the Electricity Infrastructure", International Journal of Critical Infrastructures, Volume 2, Number 4/2006, pp. 396-411.