Automated Situational Awareness Technologies for Robust and Resilient Fossil Energy Power Generation

Prepared for: 2019 Annual Project Review Meeting for Crosscutting, Rare Earth Elements, Gasification and Transformative Power Generation

About Ridgetop Group Inc.

- Industry Leader in Advanced Diagnostics and Prognostics, PHM, and CBM
- Proven track record of commercialization
- Established in 2000, headquartered in Tucson, AZ
- Design services, circuitry for harsh environments, prognostics and condition-based maintenance (CBM) solutions
- AS9100D and ISO9001:2015 Certified
- Two divisions: Semiconductor & Precision Instruments (SPI) and Advanced Diagnostics & Prognostics (ADP)
- Commercial and Government Client base
- Family of product lines with applications serving aviation, transportation, energy grid, oil/gas, communication and power system sectors.

Research and Industry Partners

TEP
Tucson Electric Power

Tucson Electric Power (TEP) has been providing power to Southern Arizona residents since 1892. Today, TEP is providing safe and reliable power to the Tucson metropolitan area with a population of approximately 1 million people and growing. TEP and its parent company UNS Energy Corporation were acquired by Fortis in August 2014. Fortis is Canada's largest investor-owned utility company that is providing gas and electric power services to more than 3 million customers in Canada, the U.S., and Caribbean.

Established in 1885, the University of Arizona is the state's land-grant university. The UA is recognized as a global leader in research, bringing more than $622 million in research investment each year, and ranking 22nd among all public universities. The UA is a member of the Association of American Universities, the 62 leading public and private research universities. It benefits the state with an estimated economic impact of $8.3 billion annually. The UA has been named a Hispanic-Serving Institution (HSI) by the US Department of Education.

Other Potential Partners in Phase II

- Georgia Power
- Southern Research
- UniSource Energy Services
- Salt River Project (SRP)
- Tri-State Generation and Transmission Association, Inc.
- Southern Company

Introduction

U.S Power Generation assets are increasingly vulnerable to malicious attacks from insider threats, terrorists, and cybercriminals. A key vulnerability to fossil fuel plants, especially with aging assets, is surreptitiously-inserted malicious intrusions caused by a lack of appropriate situational awareness by plant operators. Ridgetop's SBIR program expands the ability of plant operators to quickly locate the source of problems, whether they indicated aging equipment issues, or detection of anomalies including malicious behavior, even cyber attacks. Upon identification of problems, they can be mitigated to avoid service interruptions.

With over 8,000 fossil fuel plants in the U.S., the vulnerability of these plants to degradation and failure is a looming threat. Ridgetop's work will show that existing sensor data streams from subsystems can be mathematically combined into one single health indicator (SHI) that represents "health" of power station equipment, subsystems, and systems. Ridgetop Group Inc. (Ridgetop) has researched and developed a solution that can be applied to all sectors where real-time data, condition based maintenance (CBM), and improved data integrity are necessary.

Ridgetop, in partnership with the University of Arizona (UA), responded to SC_FOA_0001771, topic 17.a. Ridgetop had an initial plan to apply its Internet of Things (IoT) enabled smart sensors such as RotoSense and SMRT Probe to create an instrument cluster in its lab that could represent the different types of sensor readings found in a fossil fuel power plants, both hardwired and wireless. Ridgetop also created a local industry partner with Tucson Electric Power (TEP), in which they provide the entire Tucson metropolitan area with reliable utility services.

The Phase I research and development conducted in Contract DE SC-0018727, "Automated Situational Awareness Technologies for Robust and Resilient Fossil Energy Power via Multivariate Analysis", by Ridgetop and the UA Department of Systems and Industrial Engineering (SIE) centered on four concepts when determining best methods to assure data integrity within a power plant ecosystem:

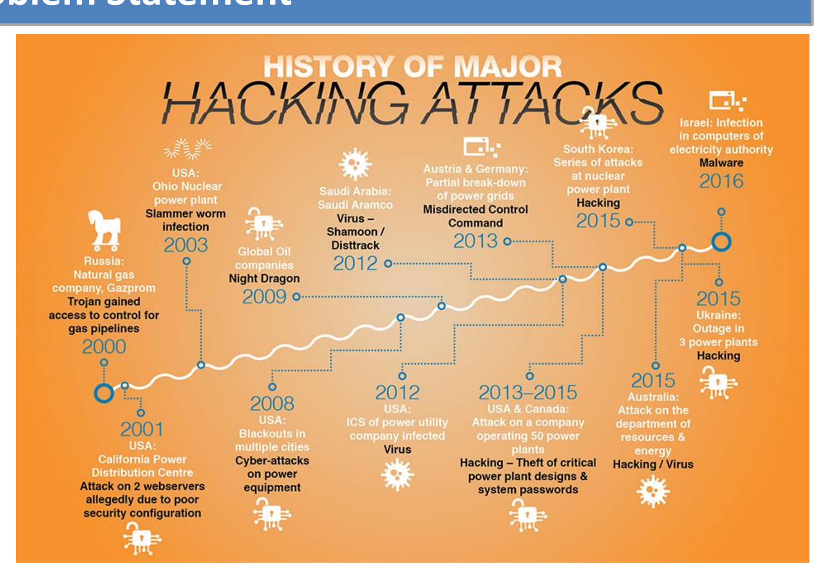
1) More efficient and more general Multivariate Analysis (MVA) algorithms for the creation of a single health indicator of components of a fossil fuel power generation.
2) Blockchain research and analysis to determine the following:
   a. Best use-case scenarios
   b. Power plant operator preference and/or need
3) Enhancing blockchain technology by combining it with MVA
4) Creation of an effective off-line simulation test bed using MATLAB Simulink

Problem Statement

Topic 17 a. Automated Situational Awareness Technologies for Robust and Resilient Fossil Energy Power Generation

The primary concerns that are identified in this research topic are a result of the vulnerabilities in U.S. power generation systems and the increasing attacks and threats from insiders and cybercriminals. Cyber attacks have grown at an alarming rate over the last two decades and it is a threat that can have serious implications if the necessary tools and security is not in place. Cybersecurity tools with automated situational awareness can trigger counter measures needed to mitigate these types of threats. Another concern of this research topic is the reliability of the fossil fuel power fleet and the data analysis to determine the health of a power plant, its subsystems and equipment.

The U.S. power infrastructure ranges from 50-year-old coal power plants to relatively new power systems that are utilizing renewable energy such as wind and solar. This diversity of equipment calls for a solution that is flexible, easy to implement, and compliant with the reliability requirements from the North American Electric Reliability Corporation (NERC).



Source: https://cdn.wartsila.com/images/default-source/twentyfour7/innovation/power-up-the-security-blanket02.jpg?sfvrsn=80169345_0

Objectives

The following list of Technical Objectives were identified in Phase I:
- Estimate state of health with secure data transfer and storage
- Evaluate and improve Multivariate Analysis (MVA)
- Construct a Simulink (Power Plant) testbed and PC-based isolated network testbed
- Develop Design of Experiments (DOE) methodology to determine prognostic thresholds for CBM, and efficacy of blockchain to sensor data protection.
- Research best methodologies for anomaly detection, failure, and malicious behavior
- Design prototype software that links Multivariate Analysis (MVA) tools with Sentinel Suite
- Integrate blockchain open ledger for multiple sensor readings with MVA methodology and determine its extensibility

Questions/Goals for SBIR 17(a)	Ridgetop's Solution
Can an effective CBM architecture be developed to monitor subsystem assets in legacy fossil fuel power plants?	Ridgetop processed multiple existing sensor values with advanced multivariate signal processing methods to extract useful and repeatable degradation signatures that can be flagged by pattern recognition algorithms.
Is multivariate analysis of sensor data streams an effective method for anomaly detection and degradation signatures?	Ridgetop has designed a set of experiments and induced faults to a test bed to gauge the efficacy of multivariate algorithmic analysis in this application.
Can new blockchain open ledger technology be used to securely log sensor readings and prevent unauthorized altering of the sensor values?	Ridgetop and the University of Arizona established two testbeds. The first was a Simulink power plant model with key functional points in the plant providing the necessary observability. The second was an isolated test network having interlinked computers that stored the sensor data with blockchain technology.
Would a vendor-independent method of monitoring and data protection for legacy power plants be useful for commercialization into a set of products?	Ridgetop worked with its commercialization partner Tech Opps and power utilities such as Tucson Electric Power, Georgia Power, and others to help identify the functionality and pricing of a product (or products) resulting from this SBIR program for financial opportunity analysis.

Methodology and Technical Approach

Multivariate Analysis Methodology

Multiple sensor measurements are difficult to analyze, therefore there is the need to create single dimensional health indicators based on the raw sensor data. The most natural way of creating such an indicator is to determine its best healthy estimate and compare it to the actual data. The distance between the data vector and its best estimate is considered the health indicator. A smaller distance indicates better health and a larger distance shows degradation, anomalies, or even intrusions. The creation of the health indicator is based on a historical data set of healthy vectors where the components of the vectors are the different sensor measurements. The space of healthy vectors is usually defined as the linear space spanned by the given healthy vectors. We look for the best estimate of any given sample vector by a linear combination of the vectors from the healthy set. The following list of MVA algorithms and methods have been developed using these principles:

A. The Multivariate state estimation technique (**MSET**) **method** is based on the least squares approach which is the same as selecting the Euclidean vector distance. This is the most commonly used approach, which has several deficiencies that Ridgetop has been able to address. One deficiency stems from the fact that the different sensor measurements are not equally important in anomaly detection; meaning the method has to be extended in order to take the importance of the different sensor measurements into account. This extension has been developed by the Ridgetop team, the software development has been performed, and the method has been tested. The second deficiency is the result of the measurement errors, when linearly dependent vectors might become independent. Therefore their spanned space might become the entire linear space, showing that any vector is healthy. Therefore, it is impossible to detect unhealthy behavior. To avoid this problem, a special **clustering algorithm** has been developed, merging close-by vectors. The third deficiency is the result of defining the healthy vector space. In this definition, any linear combination of healthy vectors is considered healthy. However, as an example, a multiple of any healthy vector by this definition is considered healthy, however the feasible range of the parameters can be violated. Therefore Ridgetop developed a special procedure where **additional constraints on the linear combination parameters** can be included into the process.

B. In selecting the **Chebyshev distance** between vectors, an alternative method is obtained. In the **Euclidean distance** the component-wise differences are squared and added. By squaring the differences we make small differences even smaller and larger differences even larger. In the case of the Chebyshev distance, these differences are either added or the weighted differences are added. Finding the best healthy estimator leads to the solution of a linear programming problem.

C. Choosing the **maximum distance** is based on selecting the largest discrepancy between the components of the best estimate and the sample vector. It shows which sensor data is the least healthy. This approach also leads to a linear programming formulation.

D. A **simple method with multivariate analysis** (MVA) has been developed to create the single dimensional health indicator (SDHI) by considering a weighted average of the individual sensor measurements where the weights are the importance factors of the sensors.

E. The **Kernel estimation method** is based on the well-known Kernel density estimation approach to find best healthy estimates.

F. The method based on the **Mahalanobis distance** makes the sample measurements independent by a linear transformation after normalizing them. The normalizing makes the average of each sensor measurement zero. Therefore the averages of the transformed vector components are also zero. The Mahalanobis distance computes the Euclidean distance of the transformed vectors from zero vector. As in the earlier procedures, smaller distance indicated better health. Selecting different vector distances, like the Chebyshev or the maximum distance, new method variants are obtained. The algorithms developed by the Ridgetop team could also consider that different sensor measurements might have different importance(s).

These methods are described in appendices A-H of the Phase I final report.

Any one or combination of the above MVA algorithms provide a single health indicator of the equipment, subsystem, or system under consideration. There are several statistics based methods to detect degradation, anomalies and even malicious behavior.

ID	Milestone	M1	M2	M3	M4	M5	M6
	Completion of Kickoff Meeting and Identification of Research Objectives						
	Completion and documentation of the Simulation Plan						
	Completion of Simulink test bed file and approval by PI						
	Best mathematical methods selected, and their most appropriate combination developed						
	Advanced hardware/software testbeds using blockchain will be constructed						
	Prototype software that links MVA tool with Sentinel Suite						
	Completed design of blockchain prototype for sensors						
	Blockchain tests are completed and documented						
	Completion of operator evaluation report						
	Completion of countermeasure report for each identified threat						
	Submitted Final Phase I report						

Figure 1.

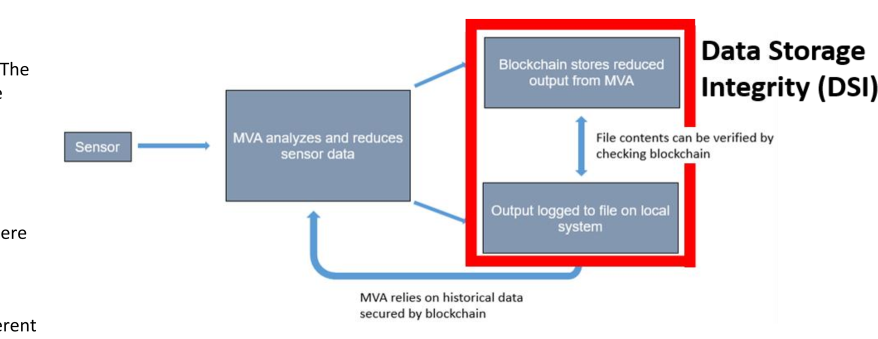


Figure 2.

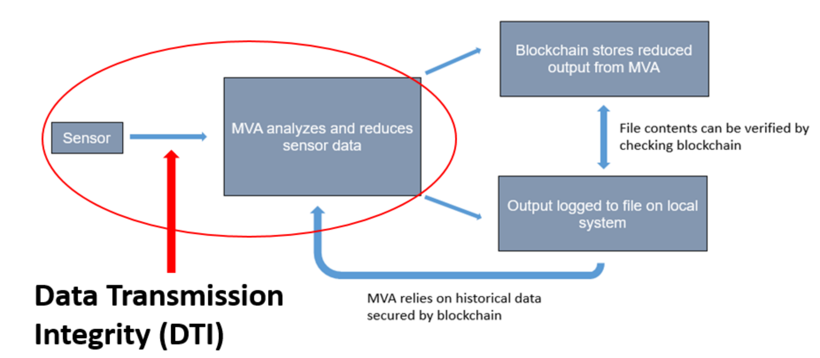


Figure 3.

Blockchain Technology

The idea of blockchain technology and how it relates to this research topic is the simple trick that the same data is transferred in multiplicity on the different channels of the data transfer system. This results in a decentralized ledger of data transactions and logged files that are immune to malicious attack on the data flow. This was verified by comparing altered / attached data files to the existing data that was previously distributed across the blockchain network. This application of the technology allows the "wrong" data to be identified and even corrected using the unaltered parallel information.

The blockchain study led by the Ridgetop research team began with identifying the most appropriate blockchain tool to use. Ridgetop researched multiple types of blockchains to find the most appropriate solution for the requested Phase I deliverables. This research presented two types of blockchains, each used for different applications.

1. Public blockchain: Anyone is allowed to join the network, see transactions, and validate transactions.
2. Private blockchain: Only certain users are allowed to join the network and interact with transactions.

Several technologies under review included Ethereum, InterPlanetary File System, a custom blockchain development, and Hyperledger Fabric. Ridgetop chose to pursue the blockchain work using Hyperledger Fabric because of its backing, development, and applications by large, trustworthy companies such as the Linux Foundation and IBM. Hyperledger Fabric also met another need of the blockchain research, which was a ***private blockchain*** rather than a ***public blockchain***. Ridgetop decided it was necessary to use a private blockchain because of the security concerns that were the primary goals that blockchain should be addressing. The proposed use-case for blockchain for power plants was to more securely hide transactions from the public (i.e. possible intruders), which made a private blockchain a requirement. Ridgetop also learned from research that blockchain is primarily used in scenarios where there are multiple organizations in the blockchain network. Since blockchain can require certain participants in each network to agree on transactions, it guarantees trust between the organizations.

Combination of Blockchain with Multivariate Analysis

The combination of blockchain technology with MVA can be done in two different ways. One is when the entire data set is sent through blockchain and processed with MVA afterwards. We prefer the other approach when MVA is used before the data is sent into the blockchain network. This approach overcomes the deficiency found in the sole application of the blockchain technology. When simultaneous sensor measurements are obtained, the shear amount of raw data can be extremely large when monitoring an entire system, subsystem, or equipment within a power plant. The MVA approach serves as a data reduction tool by condensing multiple sensor measurements into single dimensional health indicators. This has three major advantages:

1. Only single dimensional data are transferred through the blockchain communication channel, resulting in smaller computing power, small to zero transaction fees, and more efficient data transfer.
2. Equipment induced anomaly can be detected right away, which cannot be detected by the sole application of blockchain methodology.
3. A **simple algorithm for safe data transmission** can then be used in place of the blockchain approach to see if any intrusion has been attempted into the data flow. This simple algorithm is described in ***Appendix I*** of the final report. The Ridgetop team developed and verified this simple algorithm during the Phase I research project. To conduct further implementation, sensors need to be updated to have processing power. We plan to conduct further testing with the required upgraded sensors in the subsequent Phase II.

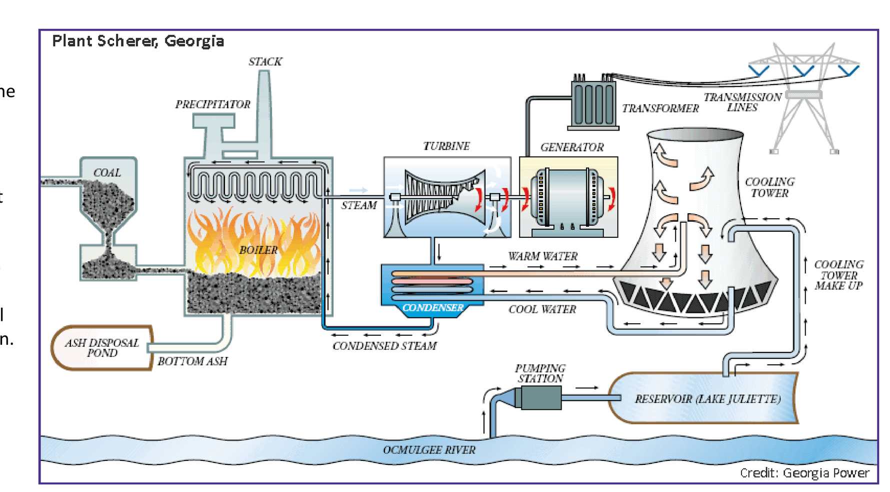


Figure 4.

Multivariate analysis and blockchain have been combined well together in two distinct ways. The first being that multivariate analysis reduces the amount of data being stored in the blockchain by retrieving a single health indicator from a set of measurements. For example, if there are **N** sensors in a system, only one measurement has to be recorded in the blockchain after multivariate analysis. This reduces the amount of data stored by a factor of **N**, which is especially relevant for blockchain as data has to be stored on each node. It also reduces the amount of data that must be transmitted to each node. In the proposed use case for this Phase 1 project, the data output by MVA has its contents stored in a blockchain network. This means that files can be checked against the blockchain network to see if they have been altered. Securing the contents of these files means that intruders are unable to alter the historical data that was recorded for machinery health. This is important because historical data is used to design models to predict future behavior and degradation. Without secure and accurate data to rely on, these models would be ineffective and possibly detrimental to companies.

MVA also relies on healthy data sets, so if an intruder successfully alters the power plant's historically logged healthy data that MVA analyzes, there could be false health indicators returned. The blockchain system that Ridgetop created protects against this vulnerability.

Ridgetop discovered 2 use cases that, along with our utility support partner, Tucson Electric Power, we envision a blockchain and MVA combined solution for the following applications:

A) **Use-case 1:** Data Storage Integrity (DSI) of logged historical sensor readings

B) **Use-case 2:** Data Transmission Integrity (DTI) from sensor to operator

Results

Evaluation Testbed and Simulation

The MVA algorithms were implemented in Python (version 3.6.3) along with a user interface to perform the analysis on data files (shown in Figure 7). To create the link from the algorithms to the blockchain, Ridgetop used the Python SDK and API for Hyperledger Fabric. The blockchain network was implemented on a private subnet of Linux based workstations as shown in Figure 5.

The University of Arizona and Ridgetop also collaborated on a Simulink model of a traditional coal power plant. One of the primary goals of the project is to monitor the sensor data to assure the operational parameters are held within the prescribed limits by MVA methods and validation by data streaming methods. A simulation model as done in Phase I is a perfect tool to achieve that goal. The Simulink model was designed according to the simulation framework for coal-fired power units.

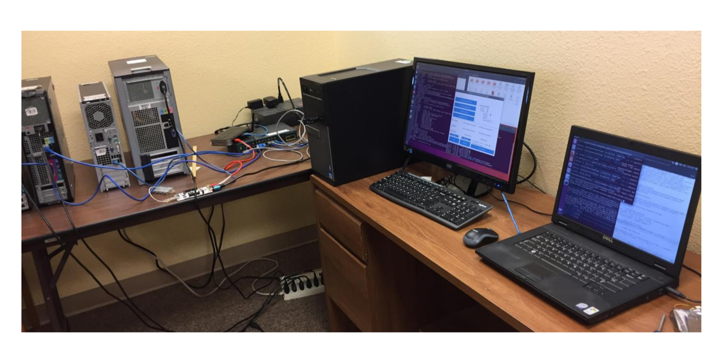


Figure 5. Ridgetop Group implementation of a private blockchain network using Hyperledger Fabric, and 5 Linux based workstations. This network was setup and evaluated at the HQ office in Tucson, Arizona.

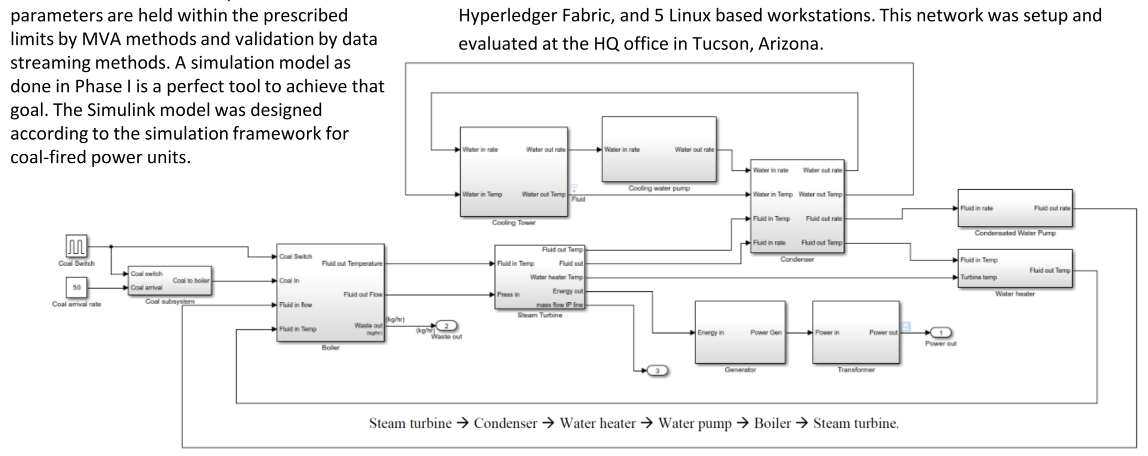


Figure 6. Simulink model produced by the University of Arizona Systems and Industrial Engineering Department. The above model is based on Figure 4 which exemplifies a typical steam-cycle coal power plant.

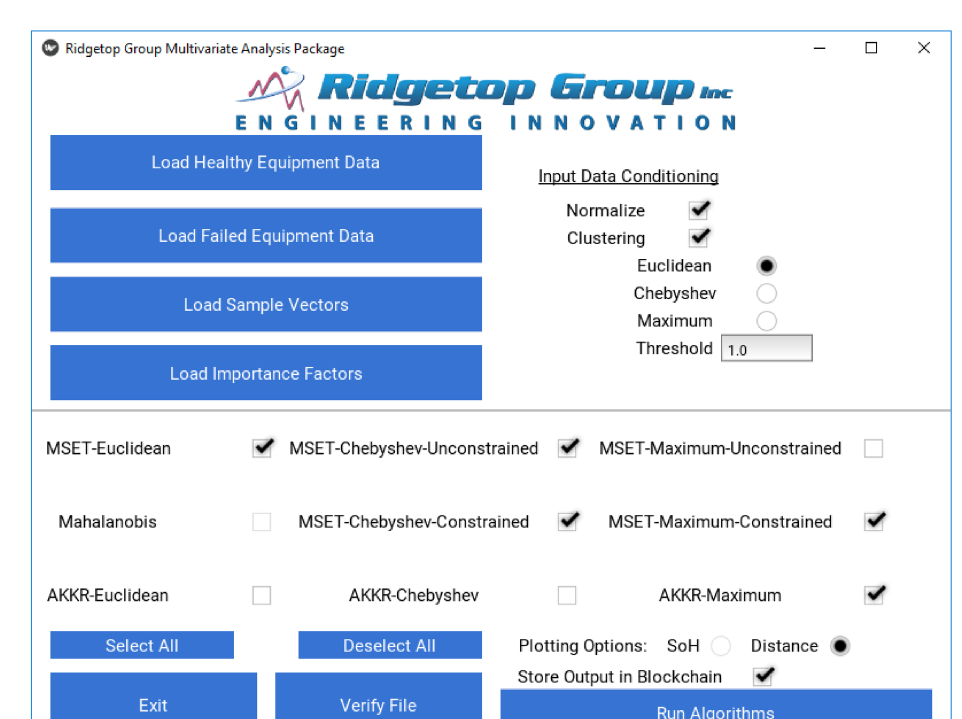


Figure 7. Prototype software program for the Multivariate Analysis Package. In Phase II Ridgetop Group intends to import this software analysis package into the Sentinel Suite cloud platform for advanced diagnostics and conditioned based maintenance.

Phase I Conclusions

Overview

During this Phase 1 project, Ridgetop researchers addressed each of the technical objectives that were identified. The Ridgetop team worked closely with the University of Arizona SIE Department, created and tested new multivariate analysis methods with blockchain technology, and established project relationships with industry partners such as TEP and Tech Opps. These accomplishments resulted in the following:

- Efficient computer algorithms that were designed to apply the MVA methodology to sensor data files.
- A blockchain network that showed proof of concept in securing logged data files by storing key-value pairs of the filename and a cryptographic hash of the file's contents.
- The blockchain implementation has been integrated so that the output from the MVA programs is securely storing the health indicators for equipment. This MVA to blockchain connection is a working verification tool that identifies malicious attempts to alter the historical health indicator data.
- A working simulation model that can be calibrated with sensor data output within a coal power plant.
- An improved understanding of the condition based maintenance practices and cybersecurity threats that are of top concerns at the TEP power plant facility in Tucson, Arizona.
- Real industry data from TEP and the Springerville, AZ power plant. Preliminary analysis has shown strong correlation between multiple types of sensors and equipment which will help calibrate the Simulink model and establish the MVA thresholds. It is planned to conduct further analysis on the data Ridgetop received from the TEP in Phase II.

Application with Industrial Partner

Ridgetop met with a team of engineers, managers, and supervisors at TEP in August 2018. The TEP team was aware of Ridgetop's project beforehand and had a number of suggestions and questions about our work. TEP was clear that the project addressed major potential issues they would want to see worked on in the future. TEP reflects a coal based power generation unit. The Ridgetop team has contacted several electric power companies, including Georgia Power, Southern Research, several agencies in Cochise County, as well as other private institutions, to cover oil/gas-based units as well as to cover the most important types of electricity generation. Near the end of the Phase I period, Ridgetop obtained a large data set from TEP and conducted some preliminary analysis along with our research partner, the UA. Further research, testing, and analysis is intended to be pursued in Phase II. Our industry partners are willing to supply their data, man-power, and other logistical assistance throughout the project.

Key Potential Customers

Potential customers from numerous agencies, companies, and programs were researched along with market predictions estimating trends and sizes from 2018 to 2023, keeping Ridgetop aware of market needs as interest increases and technology enhances. A small sampling of the current Ridgetop contact list (organization only) is shown below:

- General Electric Company
  - General Manager of Inspection Technologies
- Honeywell International, Inc.
  - NDE Manager
  - MDT Senior Engineer
  - Manager Monitoring Services
- Rockwell Automation, Inc.
- United Technologies Corporation
- General Motors

As Ridgetop continues to work with Tech Opps Consulting, Inc., as well as to expand into utilizing secondary market research specialists, we will continue to work with research agencies, private institutions, and public utilities to gather more data while expanding applications.

Target Applications in Power Plants and Fossil Energy

The equipment monitoring market is estimated to grow from USD 3.23 billion in 2018 to USD 4.47 Billion by 2023, at a CAGR of 6.7%. The market for on-premise equipment monitoring is expected to hold the largest market size during the forecast period. The monitoring solutions help organizations control all their systems and data, incurring additional cost on hardware, software, and resources employed for maintenance. The machine condition monitoring market was valued at USD 2.21 billion in 2017 and is expected to reach USD 3.50 Billion by 2024, at a CAGR of 6.7% during the forecast period. The base year considered for the study is 2017, and the forecast period is between 2018 and 2024. Ridgetop is confident that there are multiple commercialization opportunities that can benefit from the solutions and technologies that are results of this research project.

Figure 8. Example data (right image) from the Springerville, AZ power plant (left image). The tag description is the legend for 7,451 different types of data and sensor equipment.

Working Plan for Phase II

Phase I – Phase II Transition Period

Ridgetop has continued interaction with industrial partners from coal, gas, and oil based plants to validate the effectiveness of MVA in catching degrading equipment and anomalies in more settings. In addition to TEP, Ridgetop has spent time in recent weeks speaking with other energy production companies and research facilities. They have expressed an interest in partnering with Ridgetop during a Phase II period and will be able to provide additional resources in sensing equipment data and multivariate problem-solving.

Continued Research Plans

More Efficient Algorithms

Along with proving their effectiveness, Ridgetop has developed additional approaches that will be employed to increase the speed of the computer algorithms, which is critical in time sensitive power plant operations. For example, the Chebyshev and Maximum distance-based methods need the solution of linear programming problems for each sample vector. These linear programming models are identical in the coefficient matrix of the Simplex tables, only the right hand sides are changing. We plan to work out a fast algorithm to create the optimal solution of each problem based on the solutions of the previously solved problems.

Further Enhancements to Blockchain Technology

We plan to create further improvements to blockchain technology in several ways:

A) An optimal data transmission network will be constructed to reduce the size of the data flow; this is instead of transferring data from all nodes to all other nodes.

B) Implementing MVA methodology as an integral part of the blockchain by utilizing the Single Health Indicator providing a coding of the data – thereby creating hashes – since the original sensor data cannot be recovered from them.

C) A user guide for the blockchain implementations is planned to be created to show the exact steps to incorporating blockchain security in industrial applications.

System Information & Event Management (SIEM) Software Package in Sentinel Suite

In Phase I Ridgetop produced a prototype graphical user interface (GUI) with the embedded MVA algorithms. It is intended that these advanced data analytic tools will be repackaged into the Sentinel Suite software platform. Sentinel Suite is an existing Ridgetop platform that has been developed for network management, system monitoring, and is in the process of being upgraded to a complete SIEM software package that is complete with advanced data analytics and cybersecurity countermeasures. As a part of the research work plan, Ridgetop plans to investigate the promising results of game theory based network security. Ridgetop also plans to continue working with TEP, the UA, and other potential partners to apply and evaluate the software package to be created in Phase II.