Meng Yue
Brookhaven National Laboratory

Assess the Impact and Evaluate the Response to Cybersecurity Issues (AIERCI)

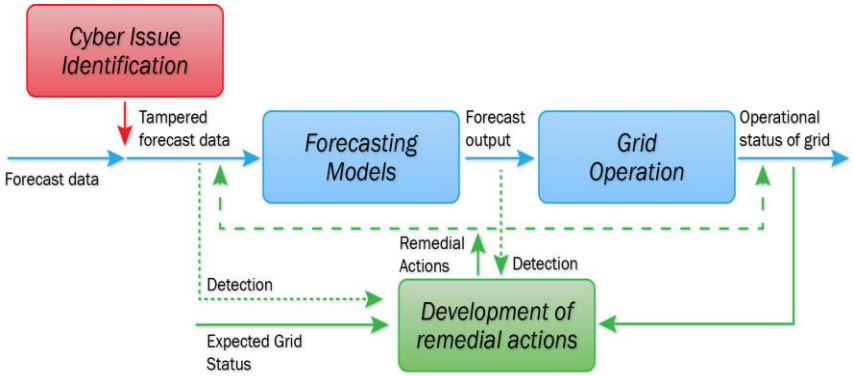
Cybersecurity for Energy Delivery Systems Peer Review

December 7-9, 2016

Summary: AIERCI Tool

Objective

- Development of an online AIERCI tool to detect, mitigate, and evaluate the impacts of cyberattacks targeting essential forecasting data that are vital to grid scheduling and operation, and vulnerable to cyberattacks.



Schedule

- Project started 02/2016 and will end in 09/2020.
- Key deliverables: An online AIERCI tool based on a cybersecure forecasting scheme and grid scheduling functions, and a demonstration by 2020.
- Will ensure the integrity of short-term forecasting data and provide data protection against cybersecurity vandalism in grid operation.

Performer:	**Brookhaven National Laboratory**
Partners:	**Argonne National Laboratory Idaho National Laboratory University of North Carolina – Charlotte Orange and Rockland Utility**
Federal Cost:	**$2.28M**
Cost Share:	**None**
Total Value of Award:	**$2.28M**
Funds Expended to Date:	**8%**

Advancing the State of the Art (SOA)

- State of the art
  - A lack of understanding of how adversaries perform cyberattacks on forecasting data and impact evaluation.
  - No online tool is available to detect, deter, evaluate and mitigate impacts of cyberattacks on forecasting data.
    - Existing handling methods to detect bad data focus on detection of point and contextual anomalies and are inadequate for coordinated cyberattacks.
- Technical approach
  - Understand and identify potential vulnerabilities and exposures in data flows between forecast model input, output, and grid operations
  - Identify the potential means or attack templates for compromising the forecasting data
  - Develop cybersecure forecasting schemes based on online anomaly detection in forecasting data induced by coordinated cyberattacks
  - Develop a planning and operational tool for assessing the impact and evaluating the response to cybersecurity issues

Advancing the State of the Art (SOA) Cont'd

- Features of the technical approach
  - The cybersecure forecasting schemes will be developed based on an integrated solution to detect various types of anomalies in forecasting data under coordinated attacks of data/models
  - An AIERCI tool will be developed by interfacing cybersecure forecasting with scheduling functions that grid operators rely on to enable an online evaluation of impacts of cyberattacks and performance of detection means
- The AIERCI Tool will be capable of dealing with compromised forecasting data and models induced by malicious attacks and evaluating performance of different detection and mitigation means.
- The AIERCI tool will be made available to utilities for handling cyber issues associated with forecasting data in both grid planning and operation.

Challenges to Success

Challenge 1: Seamless integration with energy forecasting systems and scheduling functions being used by utilities

- Solution: BNL is using systems and data that are the same as those being used by our utility partner (ORU) for realistic assessment and development

Challenge 2: Industry acceptance

- Solution: Creation of IAB for providing guidance for the tool development.
- Solution: Demonstration of the tool under an environment as close as possible to the utilities' operation

Progress to Date

Major Accomplishments

- FY 2016 Milestones
  - Creation of Advisory Board: Completed
    - Industrial Advisory Board (IAB) has been created with the following members
      - ≈ Orange and Rockland Utility (POC: Keith Brideweser)
      - ≈ EPRI (POC: Annabelle Lee)
      - ≈ New England ISO (POC: Jonathan Black)
      - ≈ SAS (Statistical Analysis Systems, POC: Jingrui Xie)
  - Identify and Justify Detection/Mitigation Methods: Completed
    - Developed new cyberattack templates that are more likely to be used by adversary
    - Developed an integrated solution and prototype implementation to better identify point, contextual, and collective anomalies while reducing the false alarm rate by combining the strength of a set of methods

Collaboration/Technology Transfer

Plans to transfer technology/knowledge to end user

- The targeted end users for the technology or knowledge include both utilities and vendors.
- To gain industry acceptance
  - Development of the tool will be performed with close consultation and input from the IAB members
  - A separate task is scheduled for a demonstration of the developed AIERCI tool
    - Testing environment of the developed tool is vital.
    - The tool demonstration will be performed in an environment as close as possible to the real operational environment
      - ≈ a control room or test bed mimicking the real-time operation based on utilities' datasets.

Next Steps for this Project

Approach for the next year or to the end of project

- Key Milestones to accomplish
  - Data flow requirements and design review: On Schedule
  - Develop short-term/very short-term cybersecure forecasting by integrating the developed anomaly detection solution: On Schedule
- Activities
  - Continue to acquire data from collaborators, perform more testing, and evaluate impacts of input compromised by cyber adversary on the forecasting data
  - Interface forecasting functions with the grid scheduling functions
  - Acquire information on collaborators' industrial control systems (ICSs) to identify potential vulnerabilities, exploit potential, and exposure in data input streams of ICS

Backup Slides

An Integrated Solution to Anomaly-based Intrusion Detection in Time Series Data for Load Forecasting

Identification of Input and Output Data in Operational Forecasting

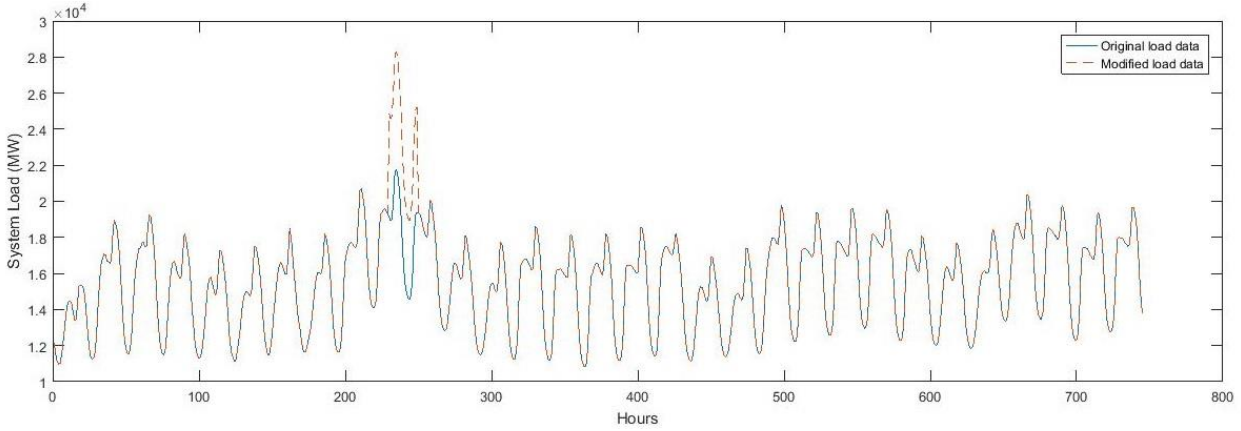
Load (Renewable Generation) Forecasting

- Input data (mainly series data)
  - Historical and forecasted weather data in the forecasting time window; the weather data may consist of temperature, humidity, dew point, irradiance, wind speed et al;
  - Historical load (renewable generation) data
    - Historical load data may consist of average load for 24-hours prior to the load forecasting time window and the same hourly load data for the previous day and week, et al.
  - Other data such as dates for holidays and public events et al.
- Output data
  - Forecasted load (renewable generation) profiles of the forecasting window.

Anomaly Types and Attack Templates for Time Series Data

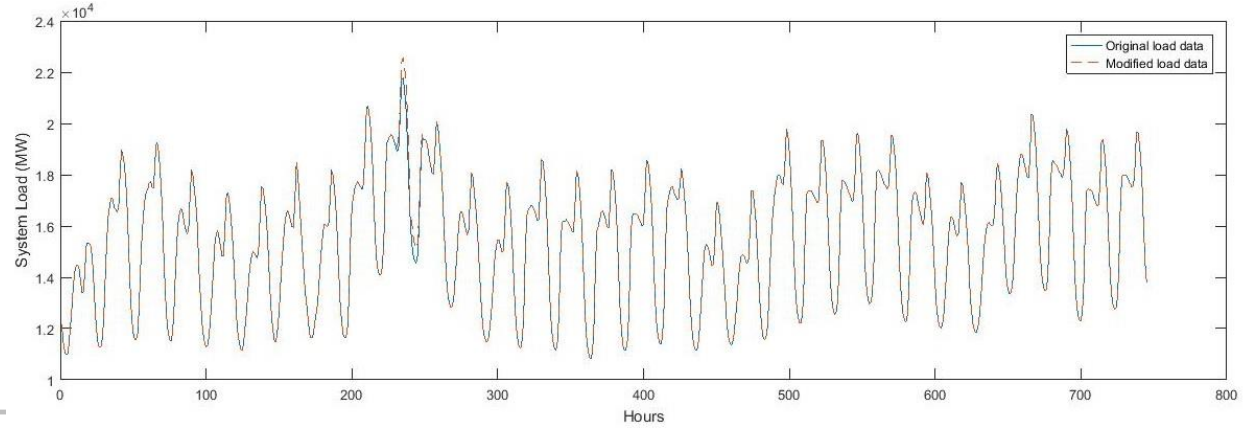
Scaling attack

- Values in the specified range will be multiplied by a constant scaling factor



Ramping attack

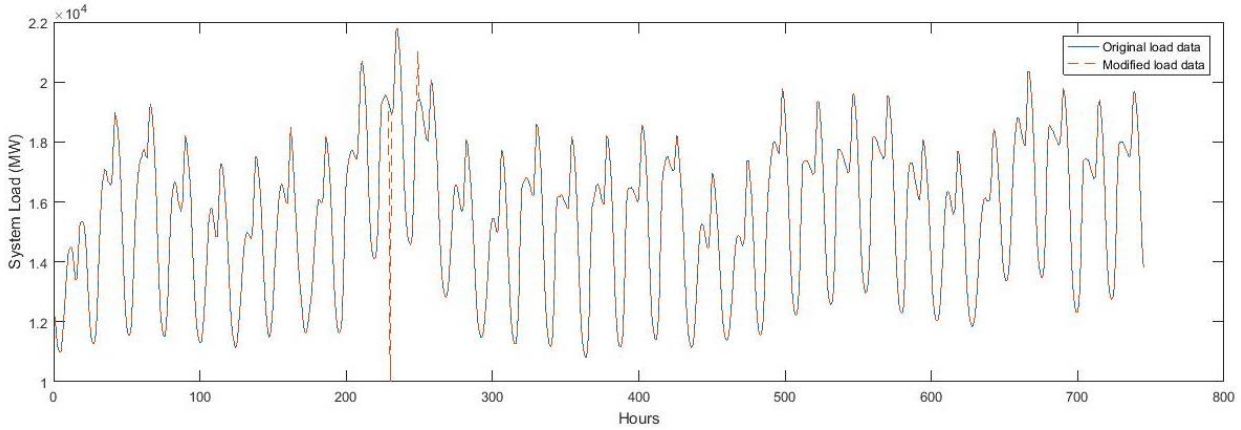
- Values in the specified range will be multiplied by a ramping function



Anomaly Types and Attack Templates for Time Series Data (cont'd)

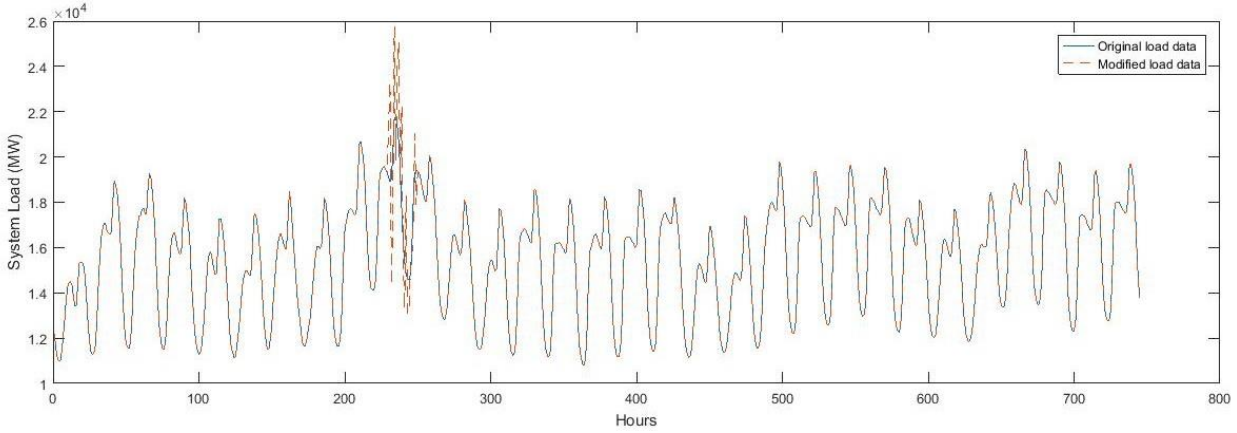
Pulse attack

- Replacing values at certain locations by individual new values



Random attack

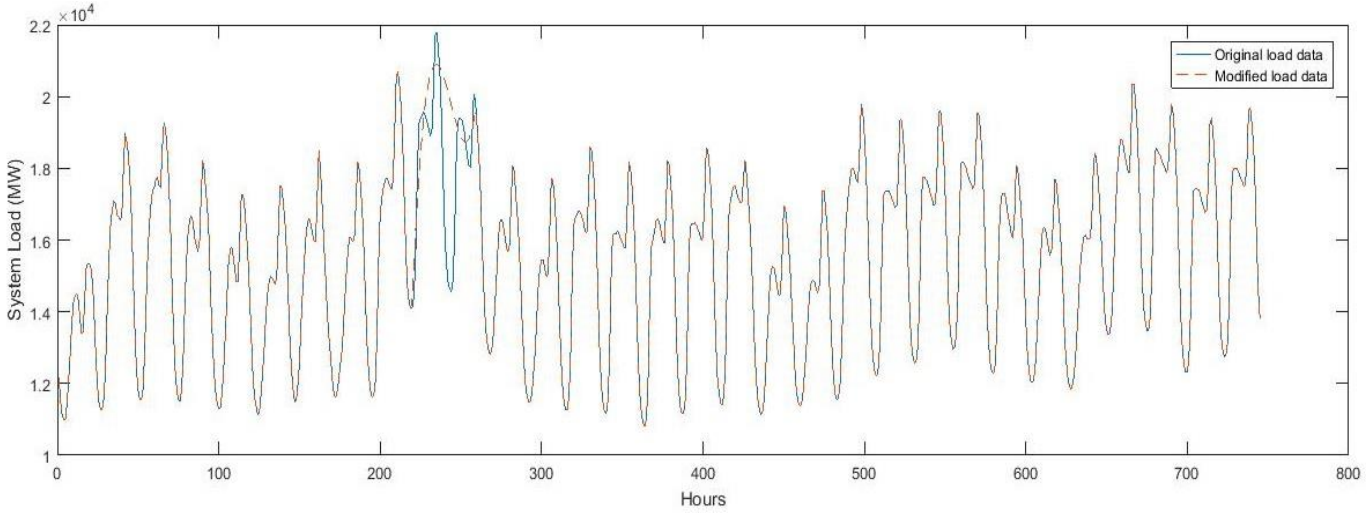
- Values in the specified range will be multiplied by a set of random numbers



Anomaly Types and Attack Templates for Time Series Data (cont'd)

Smooth-curve based attack

- Due to the nature of the previous sudden changes may be observed in the modified data points, which may be easy to detect;
- This new type of attack is performed by replacing the set of contiguous data points with a set of new values that formulate a smooth curve together with neighboring data points in the original data.



Anomaly Types and Attack Templates for Time Series Data (cont'd)

Forecasting model misuse (FMM) based attack

- This is a new type of attack that is considered an advanced technique for cyberattacks since such an attack needs a good understanding of load forecasting algorithms and models, i.e.,
  - The forecasting model will be used to generate erroneous output data using the attacked input data as input
- The basic idea for performing such attacks is that both the input data and the output data will be falsified.
  - Note that the input data are attacked using the attack templates defined in previous slides.
  - However, the modification of the output data is based on the output of the forecasting models being misused and the falsified input, and therefore, may fail certain detection methods
    - E.g., using the forecasting model and the input data to check the consistency of output data.

Anomaly Types and Attack Templates for Time Series Data (cont'd)

Forecasting model (FM) attack

- This is also a new type of attack assuming that the attackers have access and can make changes to forecasting models that are stored somewhere in the system.
  - When the forecasting models are used by the grid operator, erroneous output will be produced.
- The advantage of using this attack is that one does not need to compromise the input data to the forecasting model and still produces wrong forecasting output that is consistent with the input.
  - An example of such attacks can be changes made by the attackers to the coefficients in a regression model-based load forecasting model.

Anomaly Detection Techniques in Time Series Data

- Majority of existing anomaly detection methods are for point anomalies or contextual anomalies.
- The abnormal patterns (i.e., collective anomalies) appear to be more difficult to detect since the individual data points are not necessarily abnormal and the structure of the data has to be explored, especially in a long sequence.
- An integrated anomaly-based intrusion detection solution was developed to increase the detection rate and reduce false alarm rate by taking advantage of
  - point anomaly detection methods
  - a forecasting model based detection and
  - a unique and efficient unsupervised method for detection of abnormal pattern or discord

Anomaly Detection Techniques in Time Series Data (cont'd)

A Second Order Difference (SOD) Based Detection

- The rationale behind this approach is that the second-order differences of contiguous data points in a continuously time-invariant process should be very small or close to zero.

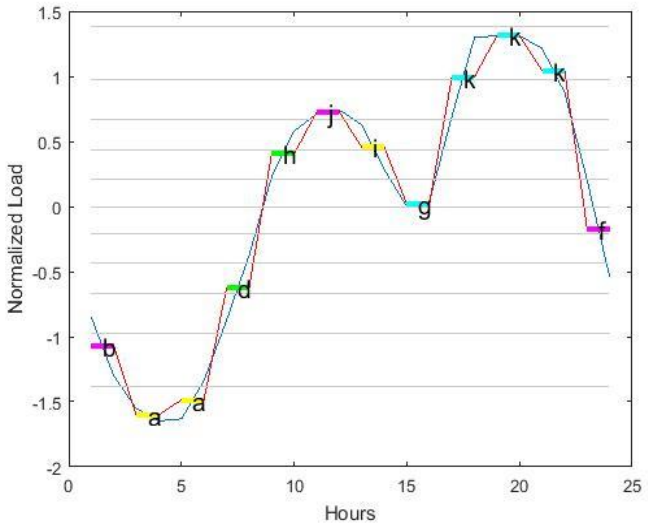
A Chebyshev Inequality (CI) Based Detection

- The property of the method is that its performance is independent of distributions of the data
- A two-staged CI-based detection method
  - The first stage attempts to determine data points that are definitely not outliers and will be used in the second stage to calculate both lower and upper bounds for detecting outliers in the original time series data.

Anomaly Detection Techniques in Time Series Data (cont'd)

A Heuristically Ordered Time (HOT) Series Based Symbolic Aggregation Approximation

- To identify "the most unusual discords or subsequences" in terms of the distances from other subsequences in a set of given time series data.
- Relies on a discretization method called "Symbolic Aggregate approXimation" (SAX) representation of time series data.
  - SAX allows both dimensionality reduction and lower bounding of L_p-norms.
  - The original time series data are represented by a piecewise aggregate approximation (PAA) of the time series.
  - A sliding window of a pre-specified width is used across the time series data to extract subsequences, obtain a PAA representation, and convert it to a SAX word.



Anomaly Detection Techniques in Time Series Data (cont'd)

A Forecasting Model Based Detection

- Take advantage of the knowledge about the correlation between input data and the output that are linked by forecasting models
  - For an offline application, the method can examine the consistency between the measurement and the calculated values using the input such that problematic data points can be further investigated.
  - Online application can be the possible detection of inconsistencies between the falsified forecasted load and the recalculated load using the forecasting model and the (normal) input data.
  - The residuals are expected to be small; therefore, the time series residuals can be evaluated using SOD- or CI-based detection techniques, as described above.
- The forecasting model based detection may also include data checking using alternative forecasting models,
  - Using alternative forecasting models will be able to identify the inconsistency among outputs from different forecasting models.

Performance Comparison between Anomaly Detection Techniques in Time Series Data

- Performance comparison was done by applying the anomaly detection methods to the sample load profiles compromised by using different types of attacks.
  - Forecasting model misuse attack also involves compromised input data and therefore, are not studied separately.
  - Forecasting model attack may be detected by using the same detection methods above but alternative models may be more efficient and therefore, is not studied separately.
- Both SOD- and CI-based methods successfully identified individual anomaly load data points introduced by scaling, ramping, pulse, and random attacks but failed to identify a smooth-curve based attack, while the HOT SAX-based method succeeded in cases.
- SOD-, CI-, and HOT SAX-based methods are all computationally efficient.

An Integrated Solution for Anomaly Detection in Time Series Forecasting Data

SOD- and CI-based methods

- Rely on thresholds that are derived from historical data and properties of new data may differ; therefore, the performance is sensitive to variations of data.
- May detect boundary anomaly points only.

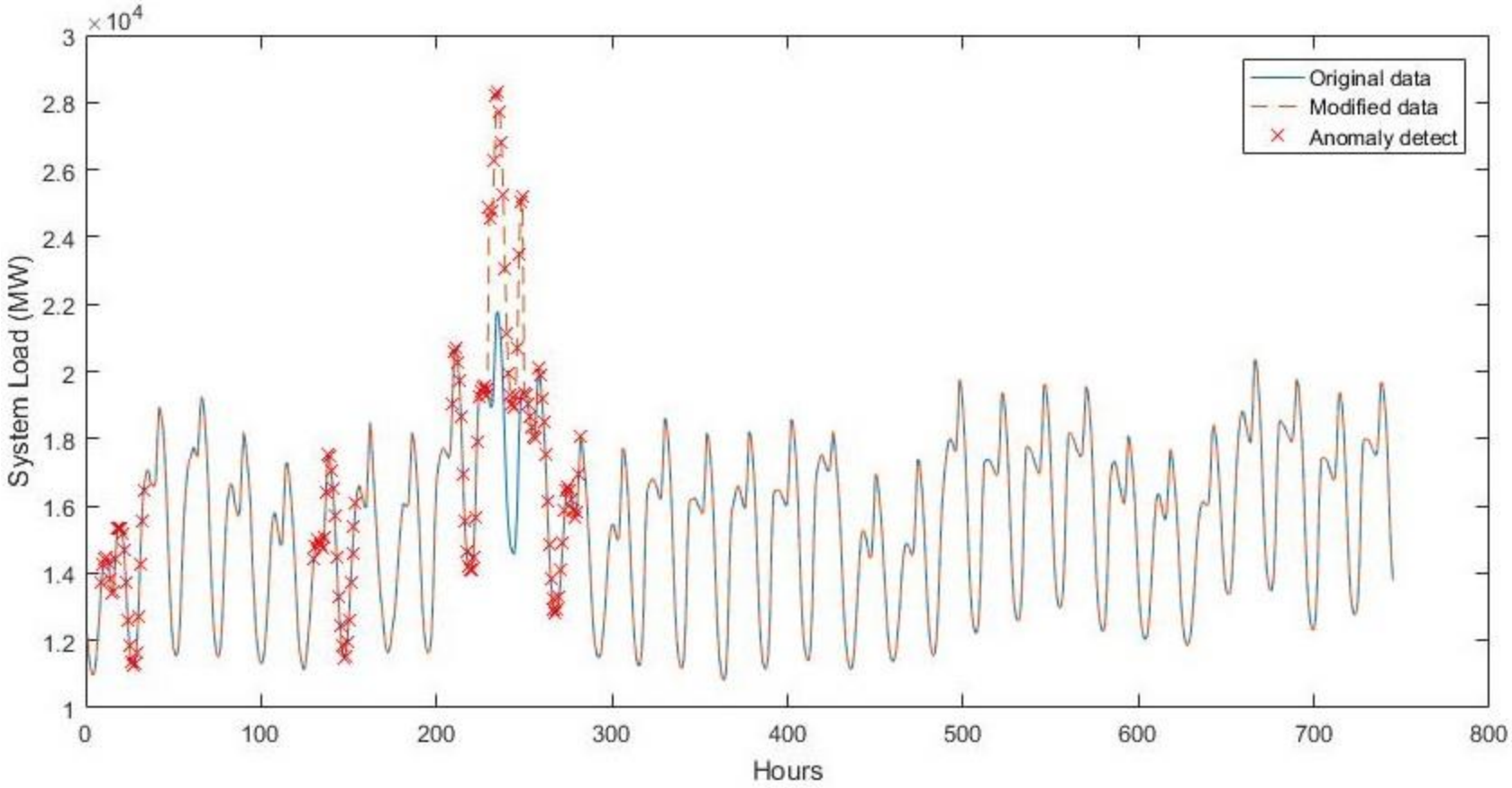
HOT SAX-based method

- Needs to specify the number of subsequences with largest distances (above a given threshold) to the rest of the subsequences, which is usually unknown and therefore, may have high false alarm rates.

Therefore, an integrated solution can be developed by combining both SOD- and CI-based methods and HOT SAX method

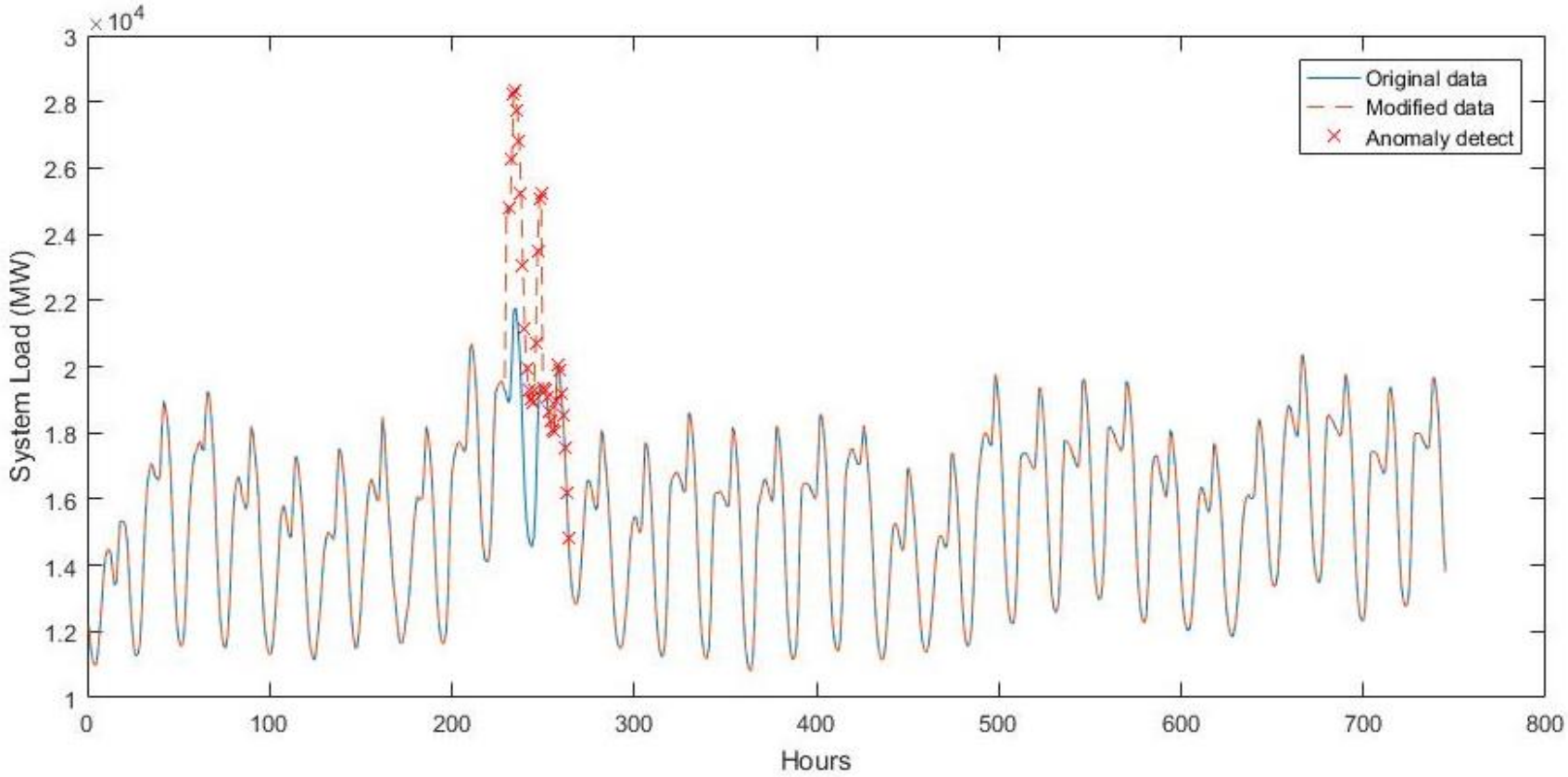
- The abnormal subsequences will be HOT SAX identified subsequences that contain/overlap with anomalies detected by SOD- and CI-based methods.
- False alarm rates can be significantly reduced while abnormal patterns or subsequences can still be captured.

An Integrated Solution for Anomaly Detection in Time Series Forecasting Data (cont'd)



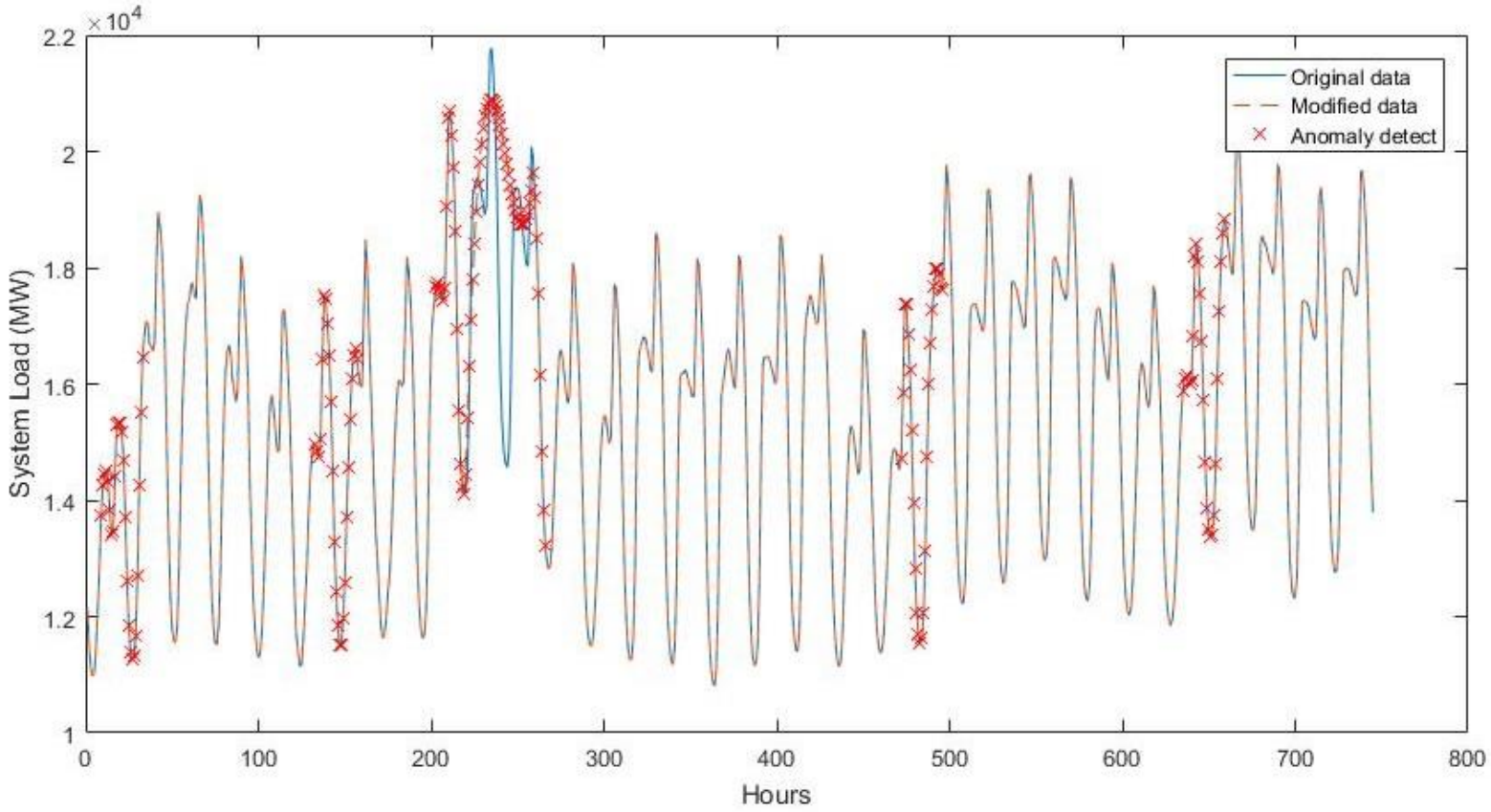
Scaling attack: Abnormal Subsequences Identified by HOT SAX-based Method only

An Integrated Solution for Anomaly Detection in Time Series Forecasting Data (cont'd)



Scaling attack: Abnormal Subsequences Identified by Integrated Solution

An Integrated Solution for Anomaly Detection in Time Series Forecasting Data (cont'd)



- Both SOD- and CI-based methods failed and the integrated solution reduces to SAX-based method only.
- There are multiple false alarms; however, the subsequence with attacked data still has the largest distance to the rest.

Mitigation Strategies

- Two major mitigation strategies are commonly available.
  - Replacing suspicious subsequences with data points from historical data on a similar day once those sequences are identified
  - Using alternative forecasting models if it is determined that the cyberattack causes a corruption of the forecasting model.
- These two mitigation strategies can be used accordingly based on the results of the integrated solution based anomaly detection.

Summary of Integrated Solution

- New attack templates that may be adopted by more sophisticated cyber intruders were developed;
- An integrated anomaly-based intrusion detection solution was developed by taking advantage of point anomaly detection and subsequence anomaly detection that are complementary to each other.
- The integrated solution takes into account the patterns of system load profiles and can effectively detect both point anomalies, contextual anomalies, and collective anomalies introduced by either bad measurements or cyberattacks.
- The integrated solution does not intend to replace but rather complement other methods such as model based detection methods.
  - The availability of multi-forecasting models can certainly help the detection of erroneous forecasting output.
- If used together with the automated implementation of common mitigation strategies, cybersecure short-term forecasting algorithms can be developed and used online as an operating tool.