ECE 445

SENIOR DESIGN LABORATORY

DESIGN DOCUMENT

Advanced Modeling and Display of International Campus Power System

---

Team #29

JIAHE LI
(jiaheli2@illinois.edu)
ERKAI YU
(erkaiyu2@illinois.edu)
TIANTONG QIAO
(tqiao4@illinois.edu)
YILANG FENG
(yilangf2@illinois.edu)

Sponsor: Prof. Ruisheng Diao

March 27, 2024

Contents

1 Introduction

1.1 Problem & Solution Overview

1.1.1 Problem Description

As an open and modern campus, Zhejiang University Haining International Campus has state-of-the-art infrastructure and cozy single dormitories. However, this also leads to a relatively high electricity consumption. In 2023, the Haining campus spent tens of millions of RMB on electricity. High electricity consumption is also common in campus around the world [1].

We found the following problems with the campus' electricity consumption:

- Lack of Awareness and Sensitivity: There is a lack of awareness and sensitivity among students and faculty regarding electricity consumption and energy conservation practices.
- Insufficient Visualization Tools: The current visualization tools for power data are not intuitive enough, hindering effective management and understanding.
- Inadequate Emergency Response Capabilities: Improved responsiveness and expanded treatment options are needed to effectively manage emergencies such as over-voltage and short circuits.

We need to promptly address these issues, as they not only hinder our sustainable energy goals but also present an opportunity to leverage the collective creativity of our campus community. By promoting conscientious energy conservation and implementing advanced visualization tools, we can pave the way towards a greener, more resilient future.

1.1.2 Solution

Our proposed solution to address the situation is to develop an Advanced Modeling and Display system for the campus power system. Specifically, we will utilize electricity consumption data from the engineering department for accurate power flow calculations. The resulting information, including current, voltage, power, and other relevant data, will be visually represented using LED strips with varying brightness and colors on a physical model.

Besides, machine learning-based algorithms such as electricity consumption forecasting and anomaly detection can be used to monitor various grid behaviors. Advanced applications like grid loss calculation and distributed wind/photovoltaic (DW/PV) power generation installation and connection are also among our considerations. In this way, we aim to visualize power distribution and usage on campus more intuitively, fostering awareness of electricity conservation among students and faculty while contributing to establishing a low-carbon modern campus.

Figure 1: Visual Aid

1.2 Visual Aid

The project's workflow, outlined in Figure 1, is a comprehensive process designed to effectively utilize collected power data. Initially, the raw power data undergoes processing using a power flow solver, which extracts relevant information vital for understanding power distribution and load balancing across the campus. This processed data is then visually represented using a solid model, offering stakeholders an intuitive understanding of the campus's power dynamics.

Furthermore, the project employs machine learning for remote data analysis. These advanced algorithms analyze historical power consumption patterns, enabling accurate forecasting and informed decision-making regarding energy management. Additionally, the project integrates power anomaly simulation techniques to simulate potential disruptions in the power supply. These simulations help stakeholders evaluate contingency plans and optimize response strategies.

1.3 High-level requirement list

- **Quantitative criteria for the front-end display.** The envisioned physical model must possess the capability to visually represent power consumption data across the entire campus accurately. It should offer detailed representations of individual buildings or substations, allowing users to discern usage patterns easily. In the case of state changes, the LEDs are expected to exhibit the desired state within a delay of **2** seconds.
- **Quantitative criteria for the machine-learning model.** The machine learning model, should be able to accurately forecast electricity consumption trends. Specifically, we expect an average MAPE of **10%** or less.
- **Quantitative criteria for the event-driven power accident simulation.** For a simulated power anomaly event, e.g., simulating a two-phase short circuit to ground in North Building A, the anomaly detection module should react and trigger an alarm within **5** seconds, and the macro-F1 score is expected to be no less than **0.95**.

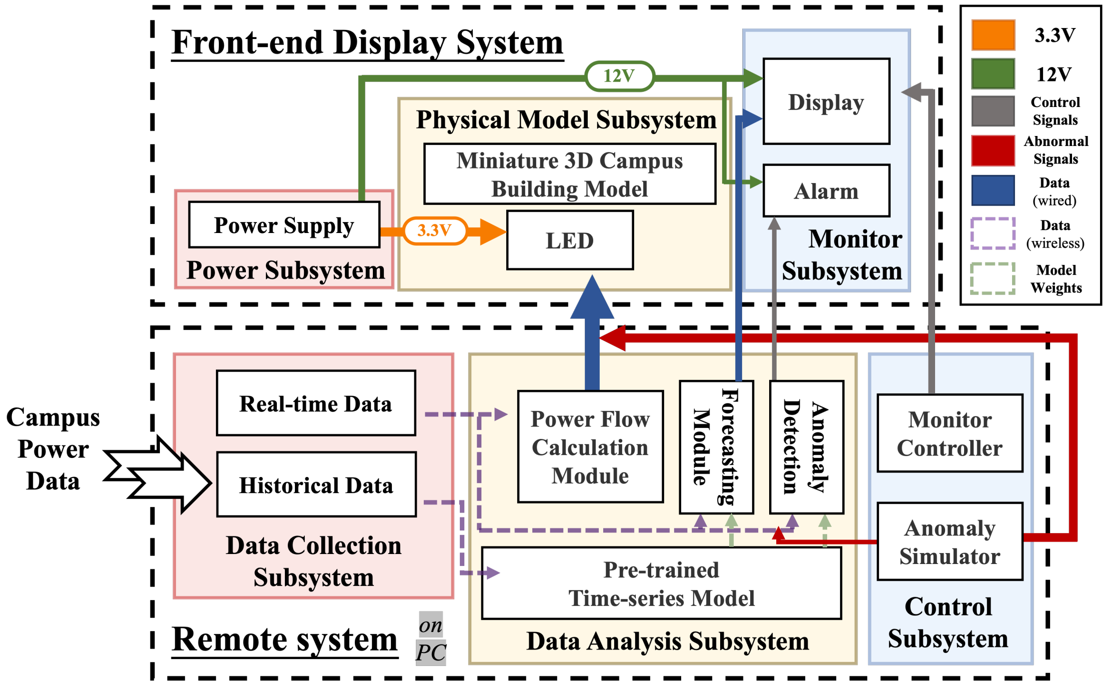

Figure 2: Block Diagram

2 Design

2.1 Block Diagram

The system (see Figure 2) first collects and analyzes real-time and historical power data from substations. The Data Analysis Subsystem processes this data to forecast usage, identify anomalies, and generate metrics like voltage and current. The Control Subsystem manages data display settings and simulates anomalies for testing. The Physical Model and Monitor Subsystems visualize the processed data, with the former using LEDs to show building usage and the latter displaying numerical data, forecasts, and anomaly alerts. The Power Subsystem provides electricity to the visualization components.

2.2 Physical Design

A rough modeling of our international campus is shown in Figure 3. The main physical objects include a sand table of each building on the International Campus, external LED strips, and interactive displays. The sand table is a square with a side length of **1** meter and the size of the display is **7** inches. Multiple external LED strips are placed next to each building to represent the different electrical data of each building. The bracket for the display is shown in the Figure 4, and the bracket is bolted to the edge of the sandbox and the display. The bracket can be adjusted up and down for different heights and left and right for different angles.

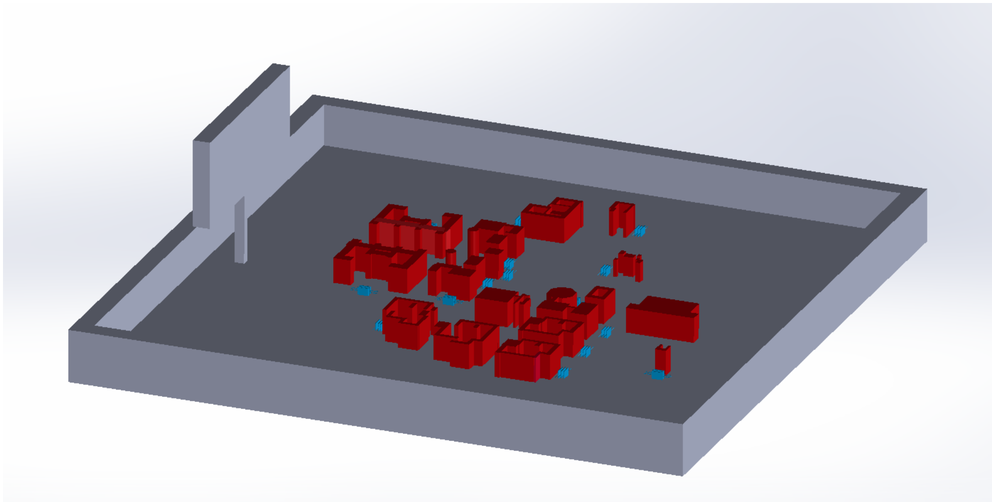

Figure 3: Sand Box of our Campus

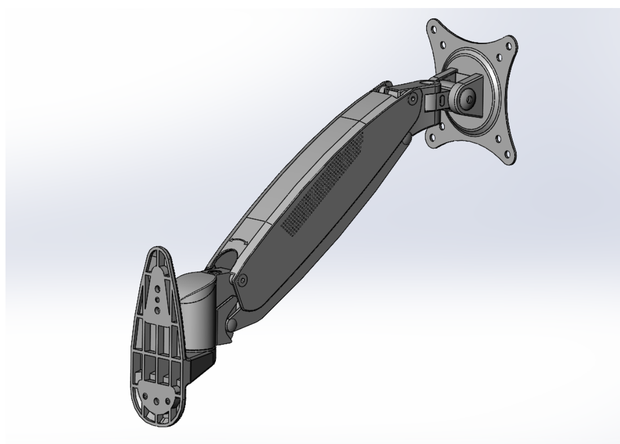

Figure 4: Screen Bracket

2.3 Remote System

2.3.1 Data Collection Subsystem

The Data Collection Subsystem collects and stores power usage data from each substation provided by the Engineering Department. This subsystem then transfers the collected data to the Data Analysis Subsystem for further analysis and utilization.

Following extensive consultations with the Support and Assurance as well as Engineering departments, we have successfully secured authorization to access real-time data housed within the database. This invaluable resource offers a granular perspective, with data points recorded at hourly intervals, detailing the power consumption in kilowatt-hours (kWh) over the preceding hour.

Furthermore, cognizant of the recurrent challenges posed by unreliable data within the distribution network, proactive measures have been undertaken. In anticipation of such scenarios, we have expanded our dataset by obtaining detailed electricity use records for the entire year 2023. These extra statistics include the aggregate energy usage within the school district, which is thoroughly documented on a monthly basis, giving us a comprehensive view of consumption trends and patterns over the course of the year.

To collect the real-time data, we have a Python script implemented to query the database in the campus. Since the data we're interested in is stored in a database and gets updated hourly, we decided to collect them once a day. The Python script runs on the database server, it will be activated at 2 a.m. each day and make queries to collect the power usage data of each building. After listing the results in files of csv format, the script will send these files to our server so that we can store the daily data and analyze it.

The Data Collection Subsystem relies on substantial power data from the Engineering Department to ensure accurate modeling. Therefore, obtaining a minimum of **1** year's worth of historical data for each substation is essential. Furthermore, a backup mechanism must be established to safeguard data integrity in the event of a system failure. Moreover, strict adherence to data privacy and security regulations is necessary to prevent any potential data breaches.

2.3.2 Data Analysis Subsystem

The Data Analysis Subsystem first processes the power consumption data obtained from the Data Collection Subsystem. Since the Data Collection Subsystem collects the electricity consumption of each substation on campus, the Data Analysis Subsystem first needs to preprocess these data into the active electricity consumption of each building. Then, the subsystem can feed the collected active power into the tidal current automatic calculation model to generate a series of desired data such as voltage and current. The data model needs to process data from all buildings simultaneously and update the model with tidal stream calculations (STEP1).

The real-time power usage data is fed into a ML model based on historical electricity consumption data, enabling the forecasting of power usage for individual buildings (STEP2) and the identification of anomalies (STEP3). The data generated by this subsystem is then transmitted to the Physical Model and Monitor subsystems for visualization purposes. For both the load forecasting model and the event detection model to satisfy high-level requirements, an acceptable degree of accuracy is required, which will be discussed in detail in subsequent sections. Also, it is imperative that the reprocessed data is transmitted to other subsystems within a maximum latency of **1** second upon receiving new data.

STEP1: Automatic power flow calculation program for distribution networks

Our project embarks on a comprehensive endeavor harnessing the power of OpenDSS, a robust open-source distribution system simulation software renowned for its efficacy in analyzing and modeling electrical power distribution networks. Leveraging OpenDSS's sophisticated cross-sectional power flow calculation and time series power flow calculation functions, we intend to delve into the intricate dynamics of our distribution network, capturing granular power data for each building at every time node. This wealth of information encompasses crucial parameters such as active power, voltage, and current, affording us unprecedented insights into the operational intricacies of our infrastructure. OpenDSS allows detailed modeling and simulation of distribution systems, enabling engineers and researchers to analyze all aspects of power flow, voltage regulation and fault analysis.

The addition of OpenDSS enriches the project by providing an advanced tool for in-depth analysis and optimization of the campus distribution infrastructure, ultimately contributing to the project's overall goal of sustainable energy management and resilience.

In the project, we need to build the campus distribution grid. The modeling and simulation results are shown in Figure 5 and 6.

Figure 5: Distribution Grid Modeling

```
CIRCUIT ELEMENT POWER FLOW

(Power Flow into element from indicated Bus)

Power Delivery Elements

  Bus     Phase     kW     +j   kvar         kVA         PF

ELEMENT = "Vsource.SOURCE"
SOURCEBUS     1     -1453.8 +j  -1717.2     2249.9       0.6461
SOURCEBUS     2     -1453.8 +j  -1717.2     2249.9       0.6461
SOURCEBUS     3     -1453.8 +j  -1717.2     2249.9       0.6461
   TERMINAL TOTAL . -4361.3 +j  -5151.6     6749.8       0.6461
SOURCEBUS     0         0.0 +j      0.0        0.0       1.0000
SOURCEBUS     0         0.0 +j      0.0        0.0       1.0000
SOURCEBUS     0         0.0 +j      0.0        0.0       1.0000
   TERMINAL TOTAL .     0.0 +j      0.0        0.0       1.0000

ELEMENT = "Line.S14"
SOURCEBUS     1      1453.8 +j   1717.2     2249.9       0.6461
SOURCEBUS     2      1453.8 +j   1717.2     2249.9       0.6461
SOURCEBUS     3      1453.8 +j   1717.2     2249.9       0.6461
   TERMINAL TOTAL .  4361.3 +j   5151.6     6749.8       0.6461
14            1     -1389.8 +j  -1397.4     1970.9       0.7052
14            2     -1389.8 +j  -1397.4     1970.9       0.7052
14            3     -1389.8 +j  -1397.4     1970.9       0.7052
   TERMINAL TOTAL . -4169.4 +j  -4192.2     5912.6       0.7052

ELEMENT = "Line.14_L5"
14            1        91.8 +j     45.2      102.3       0.8971
14            2        91.8 +j     45.2      102.3       0.8971
14            3        91.8 +j     45.2      102.3       0.8971
   TERMINAL TOTAL .   275.3 +j    135.6      306.9       0.8971
5             1       -91.6 +j    -44.4      101.8       0.9000
5             2       -91.6 +j    -44.4      101.8       0.9000
5             3       -91.6 +j    -44.4      101.8       0.9000
   TERMINAL TOTAL .  -274.8 +j   -133.1      305.3       0.9000
```

Figure 6: Power Flow Calculation

Moreover, to enhance the real-time monitoring capabilities of our system, we are poised to integrate OpenDSS seamlessly with our project framework. By tapping into the versatile API interface provided by OpenDSS, we plan to orchestrate the seamless exchange of data, enabling swift and efficient communication between our simulation environment and external applications. To this end, our team is spearheading the development of bespoke Python scripts designed to automate the construction of emulation functions, ensuring a seamless fusion of simulation and real-world dynamics. Through this amalgamation of cutting-edge technology and innovative methodologies, we endeavor to not only optimize the performance of our distribution network but also pave the way for future advancements in the realm of power system simulation and monitoring.

Overall, STEP1 provides the foundation for data analysis and presentation. A flowchart of this process is displayed in Figure 7.

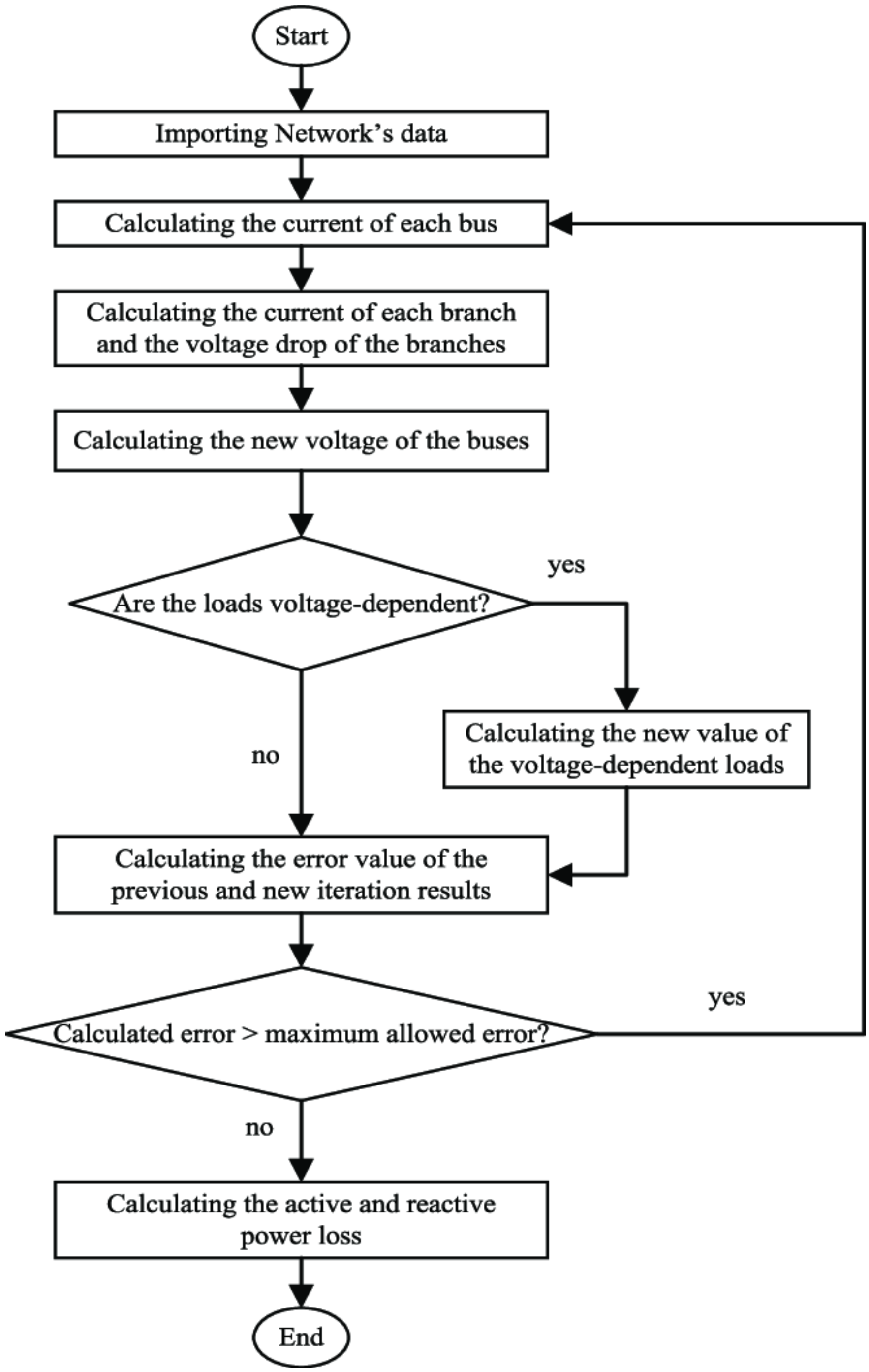

Figure 7: Flowchart of Power Flow Calculation

STEP2: Machine learning-based load forecasting

To promote energy savings on campus, it is crucial to develop accurate short-term load forecasting models for the electrical system. Specifically, machine learning methods can be employed to train models using historical and real-time electricity usage data from each individual building on campus, which can then be utilized to predict load data for future time periods. Specifically, we will predict the load for the coming day and week. The length of data used for training will be selected and adjusted according to the actual situation and error.

Based on thorough research and careful evaluation, we have selected the Long short-term memory (LSTM) [2] algorithm as our base model. This decision is driven by three key factors.

1. LSTM is a type of recurrent neural network (RNN) that is specifically designed to handle time-series data. With the ability to capture time-dependent and seasonal variations in power system load data, LSTM models are highly effective in modelling and predicting load trends. Compared with non-DL models, LSTM has the remarkable ability to autonomously acquire feature representations from input data, eliminating the need for manual feature definition and selection. This is especially beneficial for our load forecasting tasks, as load data can be affected by numerous complex factors and nonlinear relationships, making the extraction of features manually quite a challenging process.

2. LSTM models have the ability to handle noisy and outlier-ridden input data, showcasing their capability to work with imperfect input data. Techniques like dropout and regularisation can be used in LSTM models to reduce overfitting and improve generalization.

3. LSTM provides a simpler way to implement and adapt compared to the intricate models that have been prevalent in the field of time series analysis in recent times. Our requirement involves making continuous predictions on incoming real-time data, which requires the design of online prediction schemes. The simplicity of implementing LSTM and its reasonable training time make it a suitable choice for this purpose.

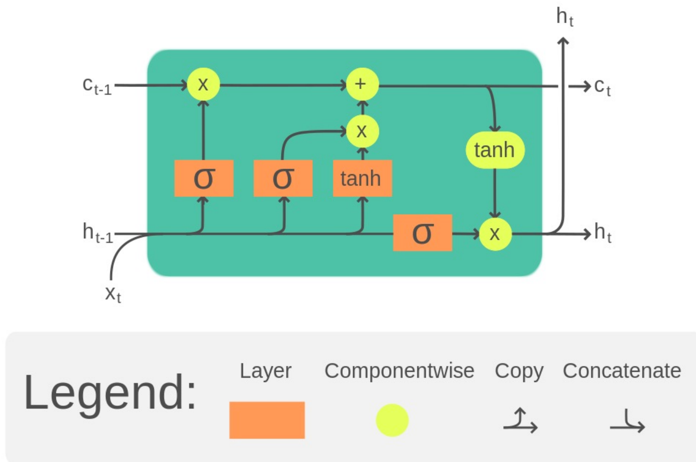

Figure 8: Flowchart of the LSTM algorithm, from Wikipedia [3]

As shown in Figure 8, LSTM uses a memory unit to store information from past time steps and decides whether to discard or retain specific information based on the current input. The memory unit consists of three gates: input, forget, and output. The input gate controls the amount of fresh information that enters the memory unit, whereas the forget gate controls how much old information is maintained. Furthermore, the output gate controls the use of information throughout the prediction process.

As per the design's functionality, the LSTM model is required to generate power consumption forecasts for the future using real-time data inputs. We propose several viable alternatives and intend to determine the approach we employ according to the specific circumstances.

- **Direct prediction based on past load data.**
- **Single-step rolling forecast.** The model predicts multiple times and uses the previous predictions as input for each subsequent prediction.
- **Multi-model single-step prediction.** A separate model is trained for each point to be predicted.
- **Multi-model rolling forecasts.** A combination of rolling forecasts and multi-model forecasts.

STEP3: Power system data-driven event detection

Anomaly detection (or event detection), on the other hand, plays a critical role in addressing the overall problem. With the detection of faults (or events) in power system, managers can promptly respond to issues, preventing potential losses and mitigating further complications.

Power system events can be categorized into major physical events and power quality phenomena. Major events, such as line trips, short circuits, generation-load imbalances,

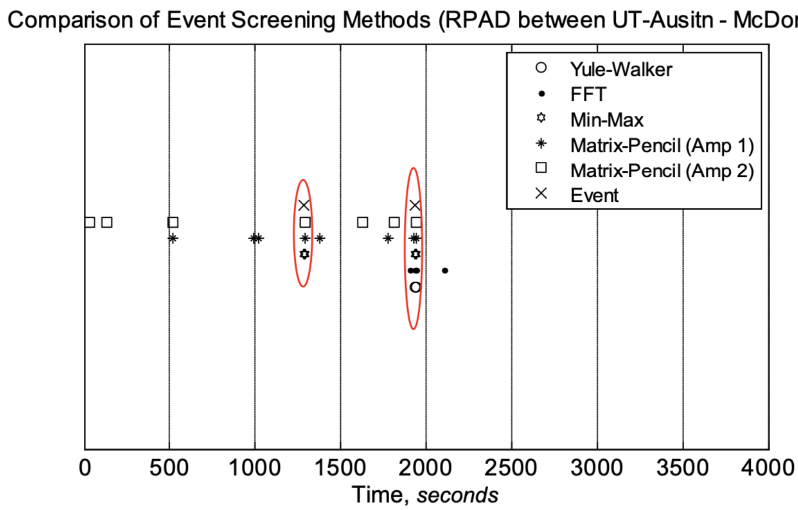

Figure 9: An example of classification results for individual features and overall classification, from [6]

equipment failures, and islanding, are directly linked to power grid components and can significantly impact the bulk power system, leading to power quality issues, large disturbances, and potential cascading failures. Power quality phenomena, on the other hand, relate to deviations in voltage, current, frequency, active/reactive power caused by minor issues like weather, contamination, equipment problems, or maintenance. Major events can also trigger severe power quality issues. These phenomena are formally defined and analyzed in IEEE power quality standards [4], [5].

It makes little sense to discuss all of the above anomalies for campus electricity usage data. Therefore, we focus on two cases, short circuit and overvoltage, and design these two events in the anomaly simulation module.

Statistical based methods can be used for anomaly detection in power usage. Specifically, the collected data is preprocessed by STEP1 to obtain time series of current and voltage for individual buildings. Statistical analysis of these time series allows anomalies to be observed. Research in related fields suggests that certain temporal or frequency domain elements can be utilised to explain the presence of events. Using these characteristics, we may develop a model that can automatically ascertain the occurrence or absence of an events.

- **Fast Fourier Transform.** It has been found that the magnitude of frequencies associated with the events in the data tends to be very high. We can choose a threshold (e.g., 3 std) above which segments are considered events.
- **Matrix-Pencil ([7]).** The matrix-pencil approach applies a sum of damped sinusoids to uniformly sampled PMU data. The damped sinusoids' parameters for fitting the PMU data are amplitude, phase angle, frequency, and damping. It is expected that if an event is present in the data, the maximum amplitude will be much bigger than that of a data frame with no events.

- **Yule-Walker Spectral.** The pyulear function can be used to calculate the power spectral density using the autoregressive Yule-Walker method. Magnitudes above a set threshold are flagged as possible events.
- **Min-Max.** We can also identify potential events by analyzing the averages and standard deviations of the differences between the maximum and minimum values within a data window. If these differences exceed a predetermined threshold, they are marked as possible events.

The 3-Sigma method can be utilized for anomaly detection in power data, such as current and voltage measurements. This method is based on the principle that normal variations in the data should fall within three standard deviations from the mean. Any data points that deviate significantly beyond this range can be considered potential anomalies. As shwon in Figrue 9, if two or more techniques identify a potential event in the same data window, the data window is tagged as containing an event.

2.3.3 Control Subsystem

The Control Subsystem enables the adjustment of data display settings on the display in the Monitor Subsystem, allowing for the selection of different time and spatial ranges. It also incorporates an Anomaly Simulator, capable of generating abnormal power consumption data to facilitate system testing and simulation of anomalies.

In order to ensure convenience, the Control Subsystem incorporates a user-friendly interface that enables manual control of the front-end display system. This control interface facilitates the selection of different time and spatial ranges to showcase various power situations. This includes historical, real-time power consumption data, power data predictions, and anomaly simulations. By transmitting control signals to the Monitor Subsystem and the Physical Model Subsystem, the Control Subsystem swiftly switches the power data display, with a response time of less than **100** milliseconds. Additionally, effective communication with other subsystems is established through standardized protocols.

2.4 Front-end Display System

2.4.1 Power Subsystem

The Power Subsystem powers the Physical Model and the Monitor subsystems, providing the proper voltage to both through a transformer.

The Power Subsystem is responsible for supplying power to the LED strips of the Physical Model Subsystem and the monitor subsystem through transformers in order to obtain different required voltages. First of all, for the light strips of the Physical Model Subsystem, the LED strips must be supplied with a continuous current of at least **500**mA and a stabilized voltage of about **3.3**V to ensure the stability of the LED strips. An overcurrent protection device is also required to prevent damage to the LED strips. For the monitor subsystem, a stabilized voltage of around **12**V must be provided for the monitor and alarm. Therefore, we use the circuit as shown in the Figure 10 to convert the DC voltage provided by the input power supply into an adjustable low voltage output by controlling the on and off state of the switching switching tubes of the PWM wave with variable input duty cycle, so as to meet the power supply requirements of different circuits. In addition, the power provision subsystem needs to be equipped with a backup power supply in case the main power supply fails. It also needs to provide status signals to indicate the health of the power supply so that it can be serviced and handled in a timely manner.

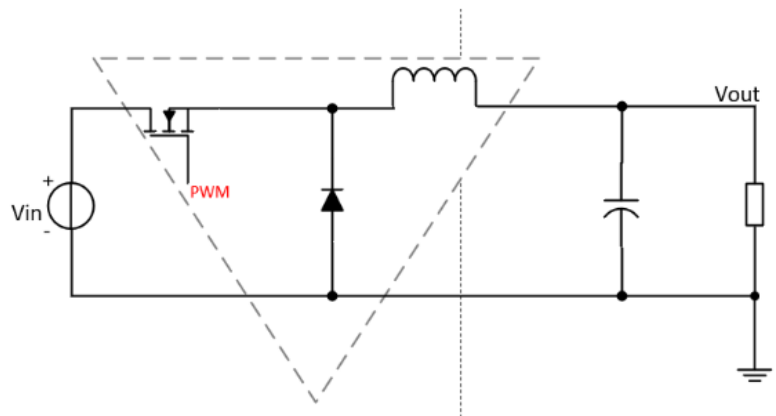

Figure 10: The Circuit of power subsystem

2.4.2 Physical Model Subsystem

The Physical Model Subsystem visualizes the real-time electrical usage of each building on campus. It receives display commands from the Control Subsystem and uses LED strips of various colors to display electrical data such as voltage and power.

The Physical Model Subsystem needs to receive display commands from the Control Subsystem within **500** milliseconds to ensure the timeliness of the visualized power data presented. The subsystem should also be equipped with at least three different colored LED strips to display voltage, power and other parameters, as shown in Figure 3. The use of LED strips to display power data allows for a more intuitive display of power data for individual buildings and zones by color, we will use green, yellow and red to represent low, medium and high power data. However, the LED strips do not directly show the exact power data values, so we need to combine this with the precise data shown in the Control Subsystem displays to get a more complete picture of the data.

Meanwhile, the campus building model needs to be as realistic as possible to facilitate real-time troubleshooting and processing. So we created a more detailed electronic model of the campus, as shown in Figure 11 and Figure 12. The Physics International Campus sandbox will be a square with **1** meter sides and a height of less than **30** centimeters. This ensures the fineness of the individual buildings and leaves enough space for the placement of the LED strips and their wiring. In addition, the Physical Model Subsystem needed to include fail-safe mechanisms to deal with issues such as LED strip failures in a timely manner.

Figure 11: Physical Model

Figure 12: Local Physical Model

2.4.3 Monitor Subsystem

To display the power usage data and make detailed analysis of it, we are going to implement a Monitor Subsystem with GUI support. The subsystem draws the hourly power usage plot of each building to help with analyzing, and it displays critical information about the building including the Electricity Consumption over the last 7 days, the System Status.

The Monitor Subsystem responds to commands from the Control Subsystem by displaying numerical real-time or historical power data of individual buildings. It can also provide estimated future values for the power data of each building. Additionally, in the event of anomaly simulation, the Monitor Subsystem can display the specific building experiencing an anomaly event, triggering an alarm accordingly.

The Monitor subsystem is required to receive display commands from the Control Subsystem within **500** milliseconds while displaying real-time power data at a refresh rate of at least once per second to ensure real-time power data. In anomaly simulation, the subsystem needs to trigger an alarm within **5** second after detecting a power failure and record the time of the failure, so as to facilitate timely handling of power failures and post-inspection. At the same time, the Monitor subsystem needs to provide a user interface for manual control and monitoring, which can be used by the observer to obtain the required or more accurate power data.

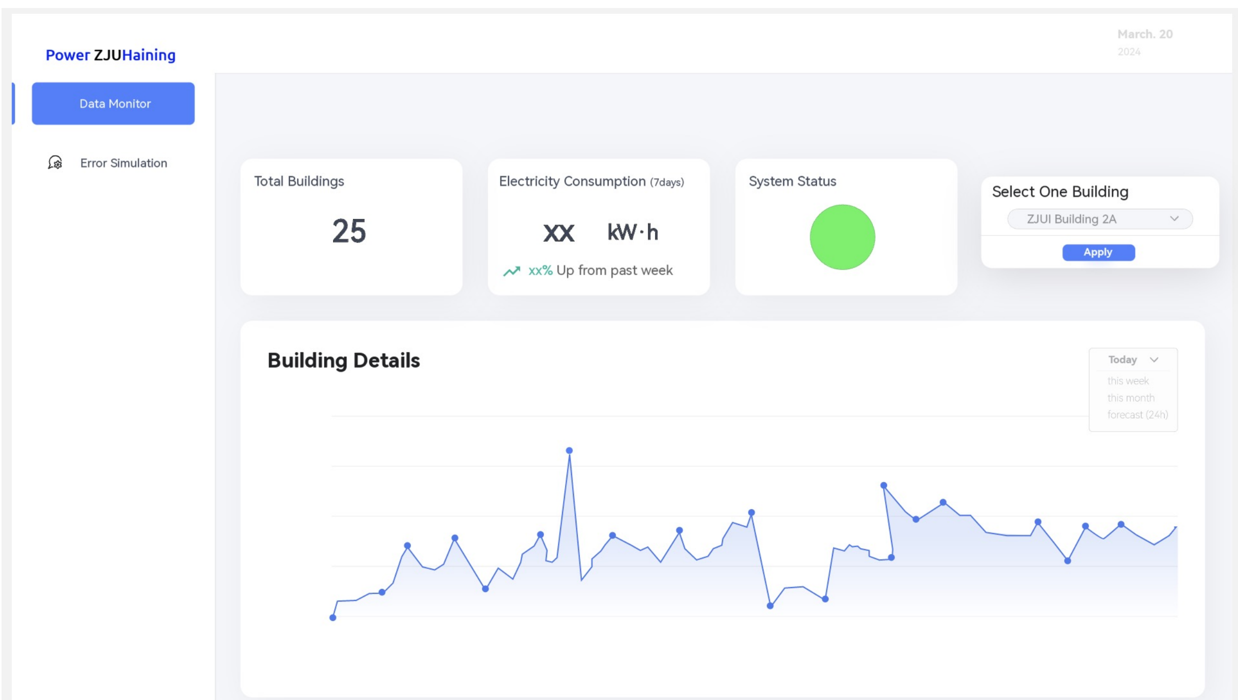

Figure 13: Monitoring System Interface Design

2.5 Requirements and Verifications table

Table 1: Overall Requirements & Verifications List

	Requirements	**Verifications**
Remote System	For power data, historical data must be accessible for at least 1 year.	The front-end display system needs to have the ability to display data from the past 1 year.
	The forecasting model must achieve a maximum MAPE of 10%	The historical data can be used as a test basis.
	The event detection model must achieve an F1 score of 0.95 or higher.	Each abnormal simulation and its neighboring normal conditions are counted from which the F1 score can be calculated.
	The remote system must maintain an update response time of $\leq$ 1s and an operation update response time of $\leq$ 100 ms	The control subsystem calculates these times, and the front-end display system shows them on the display.
Front-end Display System	The LED voltage must be 3.3V, the display voltage should be 12V. Tolerance should be within 5%.	Measure the output voltages with an oscilloscope to ensure that they remain stable.
	Color changes to the LED and display on the monitor must maintain a response time of $\leq$ 500ms	The front-end display system can calculate and show these times on the display.
	The physical model needs to include all the 26 significant buildings on campus that are electrified	Physical model will be displayed.

2.6 Tolerance Analysis

Timing Synchronization

Since power data needs to be collected and processed before display, the real-time nature of power data displayed on the front-end may be difficult to ensure. Ideally, the Data Collection Subsystem collects real-time power data at a frequency of once per second with a time tolerance of ±100 milliseconds. This includes network latency, substation reporting time variations, and system clock synchronization. The Data Analysis Subsystem processes real-time data with a tolerance of ±50 milliseconds. This includes pre-processing, model building, and providing data to other subsystems. Any delay outside of this range may affect the real-time responsiveness of the system. Also, the Control Subsystem must control the display and switching of real-time and historical data. It coordinates with the other subsystems with a time tolerance of ±20 milliseconds. This is critical for seamless switching and real-time monitoring. Exceeding this tolerance may result in inconsistent or delayed displays. In the face of possible delays, we can build mathematical models of data synchronization and processing times and use statistical analysis to determine the impact of time variations on overall system performance. Alternatively, network time protocols can be implemented to achieve accurate clock synchronization.

Power Supply Reliability

The power supply needs to provide continuous power to the LED strips of the Physical Model Subsystem and the monitor subsystem. Any failure or interruption of the power supply may result in the loss of the visual display and alarm functions. At the same time, the voltage and current supplied to multiple LED strips after passing through the transformer may be unstable, which may lead to damage or malfunction of the strips. Therefore, we need to have a backup power supply in case the main power supply fails. When a mains failure is detected, the backup power supply should be activated within 1 second, which ensures a seamless transition and avoids interruption of the LED strip operation.

Errors in the Forecasting Model of Power Usage

In accordance with the common setup of the state-of-the-art (SOTA) model in the domain under consideration [8], we choose to employ Mean Squared Error (MSE) as a metric for quantifying the dissimilarity between the predicted values and the corresponding ground truth data. Mean Squared Error is a common loss function for regression tasks, which measures the squared difference between the model's predicted output and the true values. In time series forecasting, the LSTM's output is a sequence of continuous predicted values, while the true values are the actual observations of the time series. By calculating the squared differences between the predicted sequence and the actual observations at each time step, and taking the average, we obtain the Mean Squared Error loss function.

$$\mathbf{L} = \frac{1}{M} \sum_{i=1}^{M} \left|\left|\hat{\mathbf{x}}^i - \mathbf{x}^i\right|\right|_2^2$$

Specifically, for each time step in the time series, the LSTM's output is compared to the true value, and the difference between the predicted value and the actual observation is computed. These differences are then squared, summed, and divided by the number of time steps to obtain the Mean Squared Error loss. The goal of the LSTM is to optimize this loss function, minimizing the average squared difference between the predicted sequence and the actual observations.

The final metric used to measure model performance is MAPE (Mean Absolute Percentage Error).

$$\mathbf{MAPE} = \frac{1}{M} \sum_{i=1}^{M} \left| \frac{\hat{\mathbf{x}}^i - \mathbf{x}^i}{\mathbf{x}^i} \right| \times 100$$

where $\hat{\mathbf{x}}^i$ represents the predicted value, $\mathbf{x}^i$ represents the true value, and M is the number of samples.

Errors in the Event Detection

In our design, the task of event detection in power systems can be represented as a binary classification problem. In such cases, Precision and Recall serve as commonly used metrics to assess the performance of classification models, specifically in binary classification scenarios.

Precision measures the proportion of correctly predicted positive samples out of all the samples predicted as positive. It is calculated as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

where TP represents True Positives and FP represents False Positives. A higher precision indicates a higher accuracy in predicting positive samples.

Recall measures the proportion of correctly predicted positive samples out of all the actual positive samples. It is calculated as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

where TP represents True Positives and FN represents False Negatives. A higher recall indicates a better ability to capture true positive samples.

In anomaly detection tasks, the macro-F1 score is commonly used as the primary evaluation metri. This is because anomaly detection problems typically involve highly imbalanced classes, with the normal class vastly outnumbering the anomaly class. Macro-F1 effectively captures the performance on the anomaly class, which is of greater interest, by computing the F1 score for each class independently and then taking the average. Precision and recall alone may not provide a comprehensive assessment, while micro-F1 can be skewed by the dominant normal class. Macro-F1 strikes a balance, ensuring that the model's performance on the rare but crucial anomaly class is adequately represented.

$$\text{macro-F1} = \frac{1}{2}\left(\frac{2 \times \text{precision}_p \times \text{recall}_p}{\text{precision}_p + \text{recall}_p} + \frac{2 \times \text{precision}_n \times \text{recall}_n}{\text{precision}_n + \text{recall}_n}\right)$$

3 Cost and Schedule

3.1 Cost Analysis

Table 2: Cost Analysis

Category	**Item**	**Price**
Microprocessor	RaspberryPi 4B	600RMB
Physical Model	Campus Building Model	800RMB
	Display	400RMB
	Display Bracket	300RMB
	LED*80	160RMB
Power Supply	36V DC Power Supply	300RMB
GPU	RTX 3080 Ti(Rental servers)	650RMB
Labor	4 people * 100hours * 100RMB/hour	40000RMB
Total		43210RMB

3.2 Schedule

Table 3: Schedule - 1

Date	Erkai Yu	Yilang Feng	Tiantong Qiao	Jiahe Li
3/11	Write Python scripts to make database queries and send power consumption data to local server	Construct a digital model for the physical model and made data usage requirements to verify compliance.	Confirm data types with Support and Assurance to prepare for access to the database	Researching Power Forecasting and Anomaly Detection Algorithms
3/18	Learn how to control the LEDs with single-chip microcomputers	Modeled the campus based on the physical campus landscape	Confirming the connection of each substation in the distribution network and starting modeling	Selecting alternatives for the algorithm; designing the UI for the Monitor Subsystem
3/25	Implement scripts for Raspberry Pi board to control LEDs	3D printing and sandboxing from already built 3D models and electronic models	Modeling the campus distribution network down to the substation level	Conducting selected algorithms on historical data and testing for acceptable errors
4/1	Test and debug the script for Raspberry Pi to control LED, install Raspberry Pi with LED	3D printing and sandboxing from already built 3D models and electronic models	Completing modeling of the school district's distribution network and completing testing of the offline version of the model	Selection based on data characteristics and adapting existing algorithms
4/8	Test the connection between Raspberry Pi and the data server, implement local script to receive and store data on Raspberry Pi	3D printing and sandboxing from already built 3D models and electronic models	Preparing python version of online modeling using OpenDSS API interface	Completing the code for the final time-series model and designing interfaces with other subsystems

Table 4: Schedule - 2

Date	Erkai Yu	Yilang Feng	Tiantong Qiao	Jiahe Li
4/15	Design database on Raspberry Pi to store the power data	Finish the sandbox and put LED strips around the building and connect the wiring	Completing python version of online modeling using OpenDSS API interface	Interfacing with real-time data, testing code on real-time data, checking for errors
4/22	Integrate real-time data power flow calculation on Raspberry Pi, feed it with data stored locally	Connect the wiring between the display and the sandbox so that the display can control the display state of the sandbox	Completing the online version of the real-time data power flow calculation test	Interfacing Model, Data and Monitor
4/29	Design user interaction interface with screen on Raspberry Pi	Connecting the sandbox to the siren so that the siren can give a timely alert in case of power data failure	completing power flow calculations for successful interfacing with led displays	
5/6	Integrate monitor subsystem on Raspberry Pi, help with installing the final model	Check all circuit connections, add LEDs and a circuit fault alarm system, and add a backup power supply to prevent failures	Prepare final demo and design testing cases	Prepare final demo and design testing cases

4 Ethics and Safety

Privacy. The data displayed by our system should not reflect any individual's electricity usage, as we highly value the data privacy of each individual. Thus, our system takes each building as our measuring object, to exclude any sensitive personal data while maintaining the purpose of displaying meaningful power usage data of the campus.

Social Benefits. According to IEEE Code of Ethics, we are obligated to prioritize the safety, health, and well-being of the public [9]. Furthermore, we should make diligent efforts to adhere to ethical design principles and promote sustainable development practices. Our system is designed to achieve two main goals. Firstly, it monitors the power usage of the campus to provide a safe and efficient electricity system. Secondly, it also plays a role in educating people about the value of electricity we use every day. With the model we built, we can vividly display how electricity power runs inside our campus, which urges us to use it appropriately.

Data Safety. The power usage of each building can be highly sensitive data, especially for those involving experiments. To realize our goal of power usage model display and power usage data analysis, we will preprocess the data before displaying it with our model, thus making sure that no one can reverse engineer the model to get the sensitive data. Meanwhile, the data we collected will be carefully stored to avoid any information leaks. In our project, we adhere to high standards of integrity, responsible behavior, and ethical conduct, ensuring the use of legal data sources and preventing harm to others according to [9].

Electricity Usage Safety. As our system uses a large number of LEDs to display the power consumption of the campus, it's important to monitor the functionality of the circuits and avoid potential safety issues such as fire hazards. The LEDs we use should be capable of not only long-term functioning but also smooth voltage adjustment. We will also add monitoring components for our system, in case of any unpredicted accidents.

References

[1] L. McClelland and S. W. Cook, "Energy conservation in university buildings: Encouraging and evaluating reductions in occupants' electricity use," *Evaluation Review*, vol. 4, no. 1, pp. 119–133, 1980.

[2] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[3] Wikipedia contributors, *Long short-term memory — Wikipedia, the free encyclopedia*, [Online; accessed 26-March-2024], 2024. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Long_short-term_memory&oldid=1213969171.

[4] IEEE, "IEEE Recommended Practice for Monitoring Electric Power Quality," *IEEE Std 1159-2019 (Revision of IEEE Std 1159-2009)*, pp. 1–98, 2019. DOI: 10.1109/IEEESTD.2019.8796486.

[5] IEEE, "IEEE Guide for Identifying and Improving Voltage Quality in Power Systems," *IEEE Std 1250-2018 (Revision of IEEE Std 1250-2011)*, pp. 1–, 2018. DOI: 10.1109/IEEESTD.2018.8532376.

[6] A. Allen, M. Singh, E. Muljadi, and S. Santoso, "Pmu data event detection: A user guide for power engineers," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2014.

[7] J. Quintero, G. Liu, and V. M. Venkatasubramanian, "An oscillation monitoring system for real-time detection of small-signal instability in large electric power systems," in *2007 IEEE Power Engineering Society General Meeting*, IEEE, 2007, pp. 1–8.

[8] Y. Nie, N. H. Nguyen, P. Sinthong, and J. Kalagnanam, "A time series is worth 64 words: Long-term forecasting with transformers," in *The Eleventh International Conference on Learning Representations*, 2022.

[9] *Ieee code of ethics*, Online, Visited on 03/07/2024, IEEE, 2016. [Online]. Available: https://www.ieee.org/about/corporate/governance/p7-8.html.