2024 IEEE/PES Transmission and Distribution Conference and Exposition (T&D) | 979-8-3503-1637-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/TD47997.2024.10556209

A Comparison Study of Feature Extraction and Data Fusion Techniques for Improving Cyber-Physical Situational Awareness

Logan Blakely*, Georgios Fragkos, Shamina Hossain-McKenzie, Christopher Goes, Adam Summers
Sandia National Laboratories
Albuquerque, NM, USA
*lblakel@sandia.gov

Khandaker Akramul Haque, Katherine Davis
Texas A&M University
College Station, TX, USA

***Abstract*—The power grid has historically been considered independently from the communication networks, however the physical system and the cyber system are becoming more intertwined as grid modernization initiatives push toward modern digital components. It is no longer sufficient to model the physical power system in isolation; the full cyber-physical system must be modeled for a complete system picture. Issues which were once purely cyber issues can now directly affect the physical system. This work investigates techniques for fusing cyber and physical data to analyze a scenario which includes a physical disturbance and a Denial-of-Service cyber attack which impedes control commands during the physical disturbance. Principal Component Analysis with Singular Value Decomposition, t-distributed stochastic neighbor embedding, and autoencoders are explored and compared for extracting features from cyber-only, physical-only, and cyber-physical data, qualitatively comparing the methods to provide cyber-physical situational awareness for the power system.**

***Index Terms*—feature extraction, data fusion, cyber-physical**

I. Introduction

As the electric grid modernizes and adapts to the integration of grid-edge distributed energy resource (DER) systems, data fusion across the different systems becomes increasingly important. Data fusion can be defined as the process to combine disparate data or information for understanding/estimating the state of specific quantity or system state. Thus, to understand power system states across an interconnected electric grid with DERs, it is essential to collect power system measurements across the systems and develop techniques to fuse them for understanding the power system states. State estimation (SE) is an established, highly valuable tool in electric grid operation, [1]. Nonetheless, it is no longer sufficient to only collect and process power system data, the physical system data, to assess the state of the electric grid. With the addition of smart technologies, new communication types and interfaces, and automated control functions, the grid has become a cyber-physical system in which both cyber and physical data are relevant to monitor [2], [3]. In this context, the cyber system data can be defined as network communications, host data, etc. A cyber-physical system (CPS) can be defined as a system that integrates computation, networking and physical processes that results in systems that are autonomous, intelligent, connected and collaborative [4]. To estimate CPS states, identification of relevant CPS states, CPS data types, and CPS data fusion techniques is needed. In this paper, we will focus on exploring the types of feature extraction techniques that can be applied; specifically, we explore principal component analysis (PCA) with singular value decomposition (SVD), t-distributed stochastic neighbor embedding (t-SNE), and autoencoders for for feature extraction. Using cyber-only, physical-only, and the combination of cyber-physical data the methods are assessed for their suitability to integrate the data types with temporal and structural differences.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

II. Background on Cyber-Physical Systems and Need for Fusion/CPSA

With the increasingly cyber-physical grid and rising penetration of DERs, cyber-physical situational awareness (CPSA) is needed for holistic observability into the interconnected, decentralized system. CPSA provides support to different power systems stakeholders by increasing awareness of the cyber-physical state, where state is a holistic view of whether the cyber system is compromised and if the physical system is operating in an operationally reliable state. Processes such as IEEE 1547 DER grid-support functions and communication-assisted protection schemes increase reliance on communications [5]. The highly interconnected nature of the grid with growing, distributed grid-edge presence requires greater visibility into the cyber and physical system states. It is no longer sufficient for the grid to only monitor the physical power system – the cyber-physical system must be monitored and understood to efficiently operate the evolving, cyber-physical grid as well as respond to disturbances quickly and adaptively.

Achieving CPSA in the grid is a challenging goal, especially due to the multilevel nature of the grid (e.g., establishing trust between utility, aggregator, customer levels), the lack of existent cyber-physical sensors to gather necessary concurrent data, and techniques to fuse cyber-physical data for full system CPSA. However, as more DERs and smart technologies are connected to the traditional transmission and distribution grid, it is paramount that grid operators have cyber-physical visibility into the connected system as a whole. The 2003 blackout in the northeastern U.S. demonstrated the critical need for situational awareness across utility systems; furthermore, as cyber attacks increase in frequency and sophistication, this situational awareness can no longer be limited to the physical system dynamics [6][7].

Current methods focus on single-level physical situational awareness in transmission systems. These include collecting physical system measurements (e.g., voltage, current, frequency) from PMUs and/or the supervisory control and data acquisition (SCADA) system and applying traditional state estimation techniques. Sensor data-fusion has mainly been explored for fault diagnosis, focusing on fusing physical data from disparate locations/levels. Cyber sensors are mainly used to monitor IT/enterprise networks at the utility-level such as network traffic sensors and layered monitoring frameworks with government and commercial off-the-shelf tools. Consequently, although cyber and physical data monitoring are existent in the grid, there is a significant gap for correlating the data, applying cyber-physical data-fusion techniques, and using it to obtain CPSA. In [5], the authors explore the capabilities of multi-source and multi-domain data fusion for leveraging cyber-physical data for cyber attack detection in power systems. This work focused specifically on detection of abnormalities rather than general CPSA insights that can inform planning and operation decisions in addition to response. However, we will leverage this prior work to inform approaches of fused cyber-physical data sets [6]. This paper will focus on the objective of developing novel sensor data-fusion techniques specific to CPSA insights.

III. Case Study Dataset

The dataset used in this work is an emulated version of the Western System Coordinating Council (WSCC) 9-bus model. The simulated scenario begins with a generator and line outage physical event, followed by a Denial-of-Service (DoS) attack which impedes the load-shedding signal issued by the control center. This results in an unstable system, as defined by frequency instability.

The emulation is composed of a real-time digital simulator (RTDS) that enables streaming C37.118 data from PMUs in the RTDS WSCC 9-bus model and SCEPTRE^{TM}, a Sandia industrial control system (ICS) emulation tool that enables modeling of ICS cyber/control networks and implementation of actual communication protocols such as Modbus and DNP3. The details of this emulation, scenarios, and implementation method are described in more detail in [7]. The physical disturbance data sets are collected from 8 different PMUs in the WSCC 9-bus model and the cyber disturbance data sets, roundtrip times (RTTs), are collected from 3 different relays in each of the three substations. The DoS attack targets the substation located at bus number 6. As a result of the DoS impact, the load shedding command is unable to be executed. The physical data recorded includes frequency, per-phase voltage, and per-phase current. The time resolution on the cyber data is once per second and the resolution on the PMU data is once per 33 milliseconds. For the purposes of this work, the cyber data was upsampled to the resolution of the PMU data.

IV. Data Fusion and Feature Extraction Techniques of Interest

The following sections explore four techniques of interest - PCA with SVD, t-SNE, and autoencoders. Each has advantages and disadvantages which are described in detail below. The selection of these techniques stems from their ability to perform dimensionality reduction and feature extraction high-dimensional and complex data. Each method has different characteristics, i.e., a) Autoencoders: they are neural network-based models that learn complex, non-linear representations of data that can be valuable when dealing with the complex patterns of the high-dimensional cyber-physical data. b) PCA and SVD: They are linear dimensionality reduction techniques that focus on capturing the highest variance in data. While they may not capture non-linear relationships as effectively as Autoencoders, PCA and SVD are simple and benchmarks technique to include in the analysis, and c) t-SNE: It preserves local similarities and it is well-suited for visualization tasks. One key question for data fusion is the order in which fusion and extraction steps take place. Fig. 1 illustrates two different data pipelines considering this question. In the top panel, the raw physical data and the raw cyber data are fused prior to the feature extraction step. In our case this is simply done with concatenation, i.e., one column per feature. In the bottom panel, features are extracted from the raw physical data and the raw cyber data independently and then the features are fused and used collectively. The core issue for this question is whether there are dependencies between the physical data and cyber data which can only be extracted by considering the raw data simultaneously.

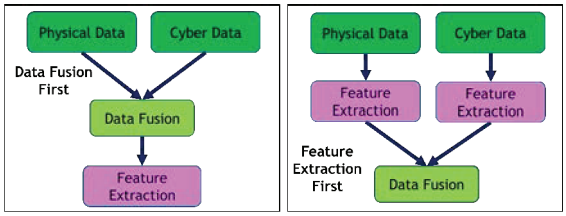

Fig. 1. Two different data pipelines for fusing physical data with cyber data

- Principal Component Analysis (PCA) is considered because it provides the best approximation of a linear model between a set of variables that may or may not be dependent on each other; PCA can be used to perform

dimensionality reduction and project the original data in a much smaller space while preserving important attributes of the data [8]. Singular Value Decomposition (SVD) is often used to do the eigen decomposition of the input matrices for a PCA analysis. For more details on SVD see, [9]

- The t-SNE algorithm is used to project a high dimensional data set into two-dimensional feature space, [10]. t-SNE preserves the ranking of small distance among the Gaussian distribution closer to zero distance and calculates the Kullback-Leibler (KL) divergence between the distribution of high dimensionality with that of low dimensionality. A perplexity parameter is used to calculate the similarity of each cluster (width of the kernel) which in turn affects the similarity scores of each point with all the other points. A t-distribution is used so that the points do not get cluttered in the middle.
- Long Short-Term Memory Autoencoder (LSTM AE) Data. The Autoencoder (AE) technique [11] is one of the most fundamental Manifold Learning techniques and it falls under the Unsupervised Learning (UL) category. The AE are Artificial Neural Networks (ANN) whose structure is symmetrical consisting of an encoder and a decoder, and this specific architecture allows them to learn an internal representation of the input data during the training process. This phase consists of reproducing the input in the output of the ANN through a series of hidden layers. As soon as the task of decreasing the input space dimensionality is completed, the internal coding must be of smaller dimensionality than the original data. The structure of the AE allows a coded representation of the input information to be obtained in the middle layer without the need for labeling or prior processing. However, in the case of the cyber-physical data, the order of daa points is important, and standard AE do not take into account temporal dependencies between data points. For this reason, Long Short-Term Memory (LSTM) AE, which can capture temporal patterns in the data, are utilized. The encoder and decoder components of the AE are built using LSTM instead of simple linear neural network layers. LSTM units are designed to memorize past units and utilize this memory to make predictions about future inputs [12].

V. Results and Analysis

This section details the results from each of the dimensionality reduction techniques described above. The results are categorized by normal versus abnormal operating conditions. That is strictly defined by the frequency in the system where frequencies greater than or equal to 61Hz or less than or equal to 59Hz are considered abnormal; this is based on operating standards from [13]. This definition, while simplistic, does provide valuable insight into the performance of the dimensionality reduction techniques and can be detailed further in the future. Voltage and other measures of system health are necessary for a complete picture and will be considered in future work, as well as taking into account the transitional period from normal to abnormal conditions which is not considered here. Future work will compare these techniques quantatively using a downstream classification task, however, we can compare these techniques implicitly through their internal metrics. In particular, for the autoencoders we can measure the reconstruction error (lower reconstruction error means better data patterns performance), for PCA the cumulative explained variance (higher variance means higher preservation of the data information) and for t-SNE there is not a direct reconstruction error but we can compute the KL divergence (higher KL divergence means better separation). Additionally, the visualizations of all the dimensionality techniques' results are provided to visually inspect how well they capture the data's underlying structure. Further testing using downstream classification of events, clustering metrics, and robustness testing is planned for future work.

A. PCA with SVD

The Matlab implementation was used for the PCA analysis, and the sklearn implementation was used for the SVD analysis. In both cases, normalization was performed on the input data prior to applying the technique. Fig. 2 shows the top two principal component vectors (blue) for each feature, with cyber-only data in a), physical-only data in b), and the combination of cyber-physical data in c). Individual samples are plotted in red using their top two components. The tables in each pane indicate the explainability of variance in that component as a percent for each of the top four principal components in each case. Circles indicate potential clusters within the PCA results. From analysis using the physical-only data we know that the first component in the physical-only data and the cyber-physical data corresponds to frequency.

The scatterplot using the just the SVD method for calculating the components is shown in Fig. 3. The results for cyber-only data are shown in a), physical-only data in b) and cyber-physical data in c). Red points indicate abnormal operating conditions, defined by system frequency, blue points indicate normal operating conditions, and green points indicate the DoS attack. For plotting purposes three SVD components are shown. In subfigure a) we see several tight groupings of datapoints. As the cyber data is not directly related to the frequency disturbance it is not surprising that the red points overlap with some of the blue points, and there is some separation for the DoS attack points (green). Subfigures b) and c) are nearly identical to each other which implies that the physical features are dominating the dimensionality reduction. This confirms the feature component explainability results discussed above. One reason for this may be the fact that there are significantly more physical features compared to cyber features, and intuitively we would expect the frequency to be the most important feature in this event. Another possibility for the similarity between the cyber-physical and physical only results in Fig. 3 is the issue of the differing time resolution between the cyber data and the physical data. These will be addressed in future work. In subplot b) and c), while

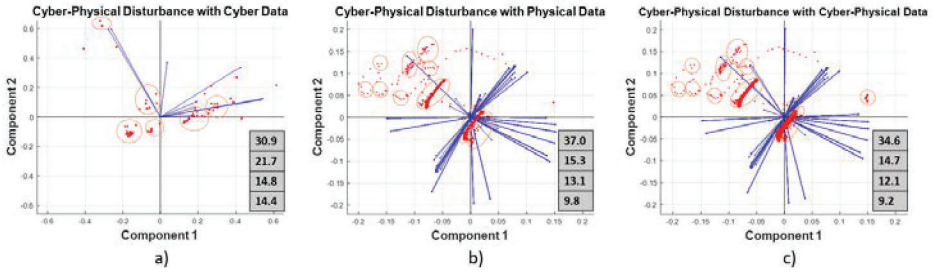

Fig. 2. PCA components for a) cyber-only, b) physical-only, and c) cyber-physical

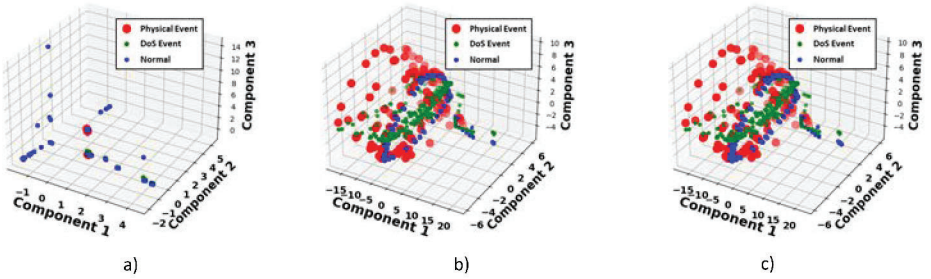

Fig. 3. SVD top 3 components for a) cyber-only, b) physical frequency only, and c) cyber-physical. Abnormal power-system operating points (as defined by frequency) shown in red and normal operating points shown in blue.

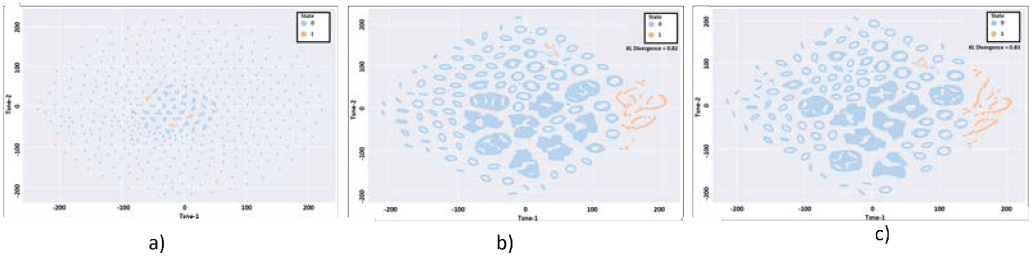

Fig. 4. t-SNE results for a) cyber-only, b) physical frequency only, and c) cyber-physical

not completely separated, we do see three distinct groupings corresponding to abnormal, normal, and DoS conditions which indicate that this method may be well-suited for use in this type of cyber-physical analysis.

B. t-SNE

The results for the t-SNE method are shown in Fig. 4. The cyber-only data is shown in a), the physical-only data in b) and the combination cyber-physical data in c). Abnormal datapoints (defined by frequency) are shown in orange and normal operating points are shown in blue. There are no distinguishing clusters in the cyber-only case, a). Again we see good clustering for physical-only and cyber-physical cases, where the results are nearly identical between the two cases.

C. Autoencoder

The results from the autoencoder testing are shown in Fig. 5. For cyber-only data, a), physical-only data, b), and the combination of cyber-physical data, c), the MSE reconstruction error for the autoencoder is shown. The normal operating conditions are shown in the top row and the abnormal operating conditions are shown in the bottom row, with the loss shown at the top of each panel. Notice that in each case the reconstruction loss for the abnormal points is significantly higher than for the normal conditions, an order of magnitude for cyber-only and cyber-physical data and two orders of magnitude for the physical-only case. One interesting finding is that for the AE method, there are some key differences in the physical-only versus cyber-physical results, unlike in the SVD case where there is no discernible difference between subplots b) and c) in Fig. 3. Here in Fig. 5 we see an increase in the loss for the normal operating condition (top row) of in the cyber-physical case, subplot c) compared to the physical-only case, subplot b), by a full two orders of magnitude. In this case, the cyber-physical data contains additional features and interactions that are not present in the physical-only data. In other words, the feature space of the cyber-physical data has more dimensions compared to the physical-only dataset, which makes it more difficult for the auto encoder. Additionally, for

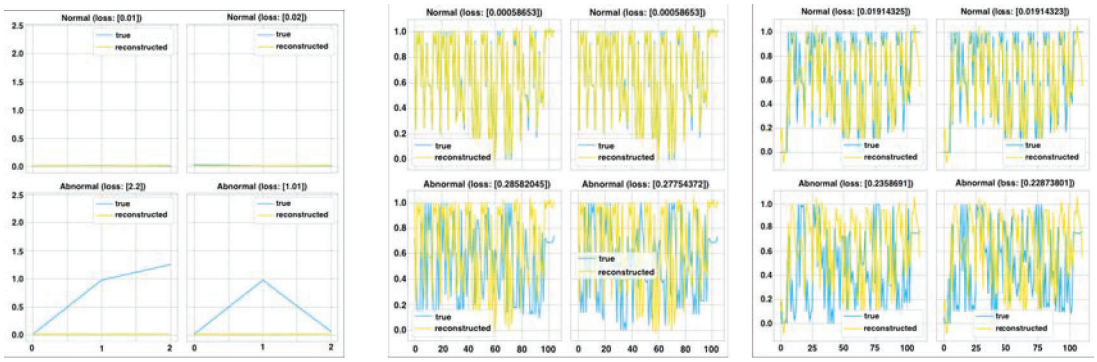

Fig. 5. Autoencoder reconstruction errors for a) cyber-only, b) physical frequency only, and c) cyber-physical. Top row shows normal operating condition (as defined by frequency) and the bottom row shows results for abnormal operating condition points. MSE loss is shown at the top of each pane

the abnormal operating case (bottom) row there is actually a small decrease, although same order of magnitude, in the loss for the cyber-physical case, subplot c) compared to the physical-only case, subplot b). This increase is likely due to the additional contextual information provided by including the cyber data (such as high RTTs during the DoS attack). Autoencoders are well placed to provide preliminary results on the classification task as well. For the case shown in Fig. 5, the autoencoder achieved 100% accuracy in distinguishing normal from abnormal operating conditions using cyber-only data, physical-only data, and the combination cyber-physical data. For these classification results abnormal conditions were defined using both the DoS event and the physical frequency disturbance. This indicates that autoencoders are extremely promising for use in this data fusion data.

VI. Conclusions and Next Steps

This work compares PCA, SVD, t-SNE, and autoencoders for the task of feature extraction on cyber and physical data for power systems applications. As the cyber and physical components of the power system become more and more intertwined, ensuring that both are considered is a critical priority. Each of the techniques was demonstrated to be a feasible choice for the data fusion and feature extraction task in cyber-physical systems. PCA with SVD demonstrated good clustering for normal versus abnormal conditions, as did the t-SNE technique. The autoencoder approach provides insight via the reconstruction errors and demonstrated excellent preliminary classification results. One differentiating factor between techniques is the higher runtime complexity of the autoencoder approach compared to PCA with SVD and t-SNE. However, the ability of autoencoders to capture non-linearity in the data may provide an advantage in feature extraction. Future work includes detailed investigation of the event classification process, a more fine-grained definition of abnormal conditions and the transition to abnormal conditions, further analysis of the impact of the differing time resolution between the cyber and physical data streams, and exploration of determining the precise timing of an event. Fusion of cyber data with physical measured data will continue to be a key component for CPSA in the modern power grid. A clear understanding of the challenges and opportunities of leveraging the cyber and physical data simultaneously will be critical.

References

[1] F. F. Wu, "Power system state estimation: a survey," *International Journal of Electrical Power & Energy Systems*, vol. 12, no. 2, pp. 80–87, 1990.

[2] N. Jacobs, S. Hossain-McKenzie, A. Summers, C. B. Jones, B. Wright, and A. Chavez, "Cyber-Physical Observability for the Electric Grid," in *2020 IEEE Texas Power and Energy Conference (TPEC)*, Feb. 2020, pp. 1–6.

[3] S. Hossain-McKenzie, N. Jacobs, A. Summers, R. Adams, C. Goes, A. Chatterjee, A. Layton, K. Davis, and H. Huang, "Towards the characterization of cyber-physical system interdependencies in the electric grid," in *2023 IEEE Power and Energy Conference at Illinois (PECI)*, march 2023, pp. 1–8.

[4] M. Törngren and P. T. Grogan, "How to Deal with the Complexity of Future Cyber-Physical Systems?" *Designs*, vol. 2, no. 4, p. 40, Dec. 2018, number: 4 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2411-9660/2/4/40

[5] A. Sahu, Z. Mao, P. Wlazlo, H. Huang, K. Davis, A. Goulart, and S. Zonouz, "Multi-source multi-domain data fusion for cyberattack detection in power systems," *IEEE Access*, vol. 9, pp. 119 118–119 138, 2021.

[6] P. Cordeiro, A. Chavez, S. Hossain-McKenzie, A. Stenger, S. Bayless, R. Clark, S. Behrendt, J. Hawkins, and K. Davis, "Considerations for secure data exchange to achieve cyber-physical situational awareness in the electric grid," in *2023 IEEE Power and Energy Conference at Illinois (PECI)*, 2023, pp. 1–7.

[7] S. Hossain-McKenzie, N. Jacobs, A. Summers, B. Kolaczkowski, C. Goes, R. Fasano, Z. Mao, L. Al Homoud, K. Davis, and T. Overbye, "Harmonized automatic relay mitigation of nefarious intentional events (harmonie) - special protection scheme (sps)." 9 2022. [Online]. Available: https://www.osti.gov/biblio/1890265

[8] J. Han, M. Kamber, and J. Pei, "Data Mining: Concepts and Techniques, 3rd Edition [Book]," iSBN: 9780123814807. [Online]. Available: https://www.oreilly.com/library/view/data-mining-concepts/9780123814791/

[9] V. Klema and A. Laub, "The singular value decomposition: Its computation and some applications," *IEEE Transactions on Automatic Control*, vol. 25, no. 2, 1980.

[10] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne." *Journal of machine learning research*, vol. 9, no. 11, 2008.

[11] P. Baldi, "Autoencoders, unsupervised learning, and deep architectures," in *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, 2012.

[12] Y. Yu, X. Si, and J. Zhang, "A review of recurrent neural networks: Lstm cells and network architectures," *Neural Computation*, vol. 31, no. 7, 2019.

[13] N. R. Subcommitee, "Balancing and Frequency Control," North American Reliability Corporation (NERC), Tech. Rep., 01 2011.