Le Xie

data technology

the new normal

I WAS SITTING AT THE FIRST IEEE Workshop on Utility Big Data, held in San Antonio in September 2017. It was right after the historical Hurricane Harvey had devastated Texas. Valentine Emesih, CenterPoint Energy's Division vice president of grid and market operations, talked about painting the layer of real-time outage data in geographic information systems (GISs) to make customers aware how power restoration was progressing. As bad as the hurricane was, big data offered a glimpse of optimism and hope for the postemergency recovery. I could not help but think about the potential of big data and analytics in transforming the power industry.

Today, there are two major categories of experts in utilities and system operators: the IT (information technology) folks and the OT (operational technology) folks. While listening to the talks, I kept thinking that, in the future, the third pillar of the power industry would very likely be experts in "DT" (data and analytics technology) as the new normal.

Given the critical nature of the electric power infrastructure, state-of-the-art technologies in sensing, communication, and computation have been always tightly coupled with the development of the grid. In the past decade or so, thanks to the smart grid initiatives around the world, billions of dollars' worth of information and communication technologies have been deployed. They generate large amount of data in various timescales.

The industry has started to take advantage of these emerging data by addressing many problems through a data-driven approach. In addition to the visualization of data using GISs, there are many other applications. For example, how can smart-meter data be used to better calibrate nontechnical losses? How can distribution system models and measurements be used to calibrate parameters for equipment such as transducers? How could one consolidate many events in an integrated data modeling framework? How can one model and trace electrical changes at the end-user level while still keeping the complexity of the model relatively easy to handle at the transmission level? Additional use cases include storm damage prediction, high-risk feeder evaluation, and many others. As a power engineer, this opens up so many new opportunities to solve big problems with immense societal impact.

From a research point of view, the integration of data from massively deployed sensors with underlying physical models, cybercalculations, and human-based operations provides a major opportunity for the power system analysis community. At the core of this unprecedented architectural change lies the scientific challenge of modeling and operating the increasingly complex electric power system. The technical performance and economic value proposition of distributed resources could be aligned in restructured electricity markets. State-of-the-art data science can be integrated with real-time dynamic power systems and offer streaming analytics for the operation and planning of the underlying power grid. Streaming synchrophasor data can be used for the online monitoring of system anomalies and even provide classifications of root cause system events for predictive and preventive control means. Similarly, at a different timescale, the behavior of consumers subject to real-time prices can be collected and analyzed through a dynamic systems lens. A recent study suggests that the behavior could be quite nonlinear with respect to price, due to the recollection of previous price patterns.

The value of the data is likely to provide the power industry with a renewed paradigm that leads to a renaissance of its business models. Over the past decade, the power industry has been undergoing fundamental paradigm shifts due to a number of regulatory and economic drivers. The first major driver is the drastic change of the generation portfolio. Some 200 GW of renewable variable energy capacity contributes about 20% of the entire generation capacity in the United States, while 16 GW of coal generation was retired in 2015 alone, with more retirements to come in the coming decades. The reliability and security of the grid will

(continued on p. 104)

Digital Object Identifier 10.1109/MPE.2018.2793479
Date of publication: 18 April 2018

need to be carefully engineered for such changes in the generation portfolio.

The second major driver is the distribution grid's increasing level of innovation and utilization when compared to the wholesale-level transmission grid. For example, in the European Union, more than 90% of the newly installed solar photovoltaics was integrated at the lower voltage distribution systems in 2016. This prompts fundamental questions for utilities of the future. What would be the new business model that would align various stakeholders' interests under the new environment? How should the utility companies redefine their business models so that their stakeholders (including utility customers, distributed resources providers, infrastructure owners, system operators, and vendors) can be properly incentivized to improve social welfare?

In this quest, data would likely play an important role. Many value-added services from utility companies could very well become new avenues of innovations and business opportunities. Examples include home-energy efficiency advising, peak-time rebate, or even transportation electrification, just to name a few. All these would require the third pillar of the future power industry, the DT-savvy experts.

Just as data are transforming the way people conduct retail businesses, hotel businesses, and many other businesses, data are likely to have a major impact on the modernization of the electric grid. In addition to technical issues, there are policy barriers that regulate access to and the use of data. Unless there is a clear policy framework that lays out what is accessible, sharable, and securable, it would be difficult for technologies to succeed by themselves. The industry needs a clear regulatory framework to address the seams issue across organizational borders and define the ownership and usage of data in different instances.

Another key issue is how to keep the data closely coupled with the physical models. Figure 1 shows the possible vision of integrating data with underlying physical models. It would not be wise to throw away all of the wonderful knowledge that engineers and researchers have accumulated over decades and start a brand new data-only approach to power and energy systems. Instead, such models and physical principles offer tremendous insights to guide the learning aspects of data-enabled discovery. In terms of practical implication, it would be important to place data scientists and power engineers side by side so they can communicate with each other and appreciate each other's value and insights. On the research front, it would be extremely fruitful to combine machine learning and data science tools, with a deep integration of the dynamic physical model at various timescales.

The third issue is data availability and accessibility. While not every piece of data is sharable in the public domain, there are many venues that

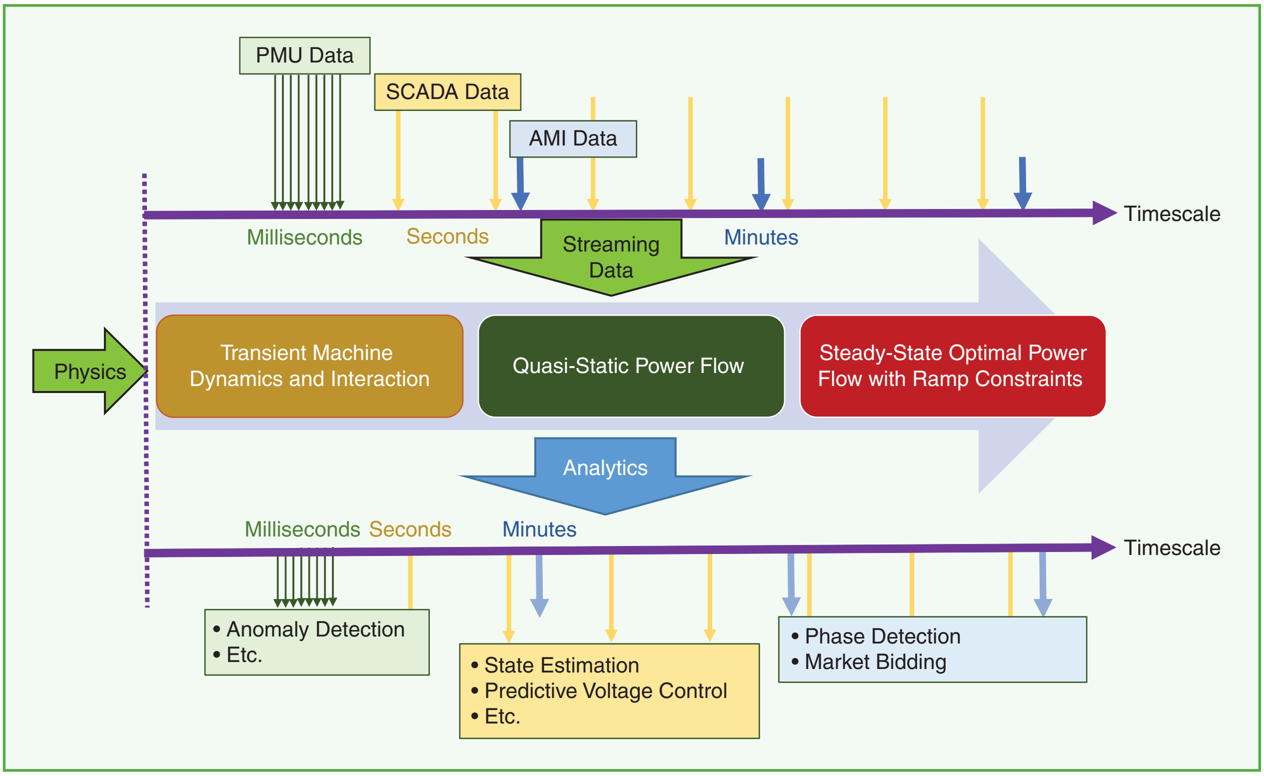

figure 1. The integration of data and physics-based models in power system operations at various timescales.

deposit actual as well as synthetic data. In the United States, for example, sharing real grid data is limited by restrictions pertaining to critical energy infrastructure information, as defined by the Patriot Act of 2001. Therefore, it would become critical for the research community to build realistic scale yet synthetic data to allow for testing of new algorithms and ideas.

The IEEE Power & Energy Society (PES) community recognized the challenges and opportunities in big data as soon as the term "big data" began to pick up momentum. After an initial proposal in July 2012 (during the PES General Meeting in San Diego), the community decided to establish a task force to study the issue of big data and analytics for grid operations; the subject is now under the newly restructured Analytical Methods for Power Systems Committee. The committee has established several working groups and task forces, focusing on data access, distribution systems analytics, and an IEEE-wide webinar series. The group has just successfully organized the first IEEE Utility Big Data Workshop in San Antonio, with more than 120 attendees from around the world. An education monograph is under development to introduce data sciences to PES. Cumulatively, ten panels and more than 1,000 participants have been involved in the activities organized by this group. This committee participated in the *IEEE Transactions on Smart Grid* special issue on big data, published in September 2016. There are several technical reports under development with the volunteers from this committee. The group welcomes more volunteers.

Overall, I am extremely optimistic about the possibilities that data may offer to power and energy systems. This could become as important to the power industry as the introduction of the digital computer to the power sector, which transformed the entire electric power industry. The introduction of data and data intelligence might transform our industry even further. We are lucky to be at a very exciting moment in history.

For Further Reading

Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

M. Wu and L. Xie, "Online detection of low-quality synchrophasor measurements: A data driven approach," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2817–2827, July 2017.

X. Geng and L. Xie, "Learning the LMP-load coupling from data: A support vector machine based approach," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1127–1138, Mar. 2017.

T. Hong, C. Chen, J. Huang, N. Lu, L. Xie, and H. Zareipour, "Guest editorial big data analytics for grid modernization," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2395–2396, Sept. 2016.

Digital Object Identifier 10.1109/MPE.2018.2816863