Data Challenges and Analytical System Requirements for Distribution Grids with High Solar Penetration

Bilal Saleem, Muhao Guo, Yang Weng

Arizona State University

Introduction to Data Challenges

Analytical System Requirements

Examples and Solutions

Conclusion

Fossil Fuels Reserves → Deplete Soon

Less Maintenance Required

- Around 30 years life of solar panels

Clean Energy

Reference:

Bhattarai, Bishnu P., et al. "Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions." IET Smart Grid 2.2 (2019): 141-154.

Lopes, João Abel Peças, et al. "The future of power systems: Challenges, trends, and upcoming paradigms." Wiley Interdisciplinary Reviews: Energy and Environment 9.3 (2020): e368.

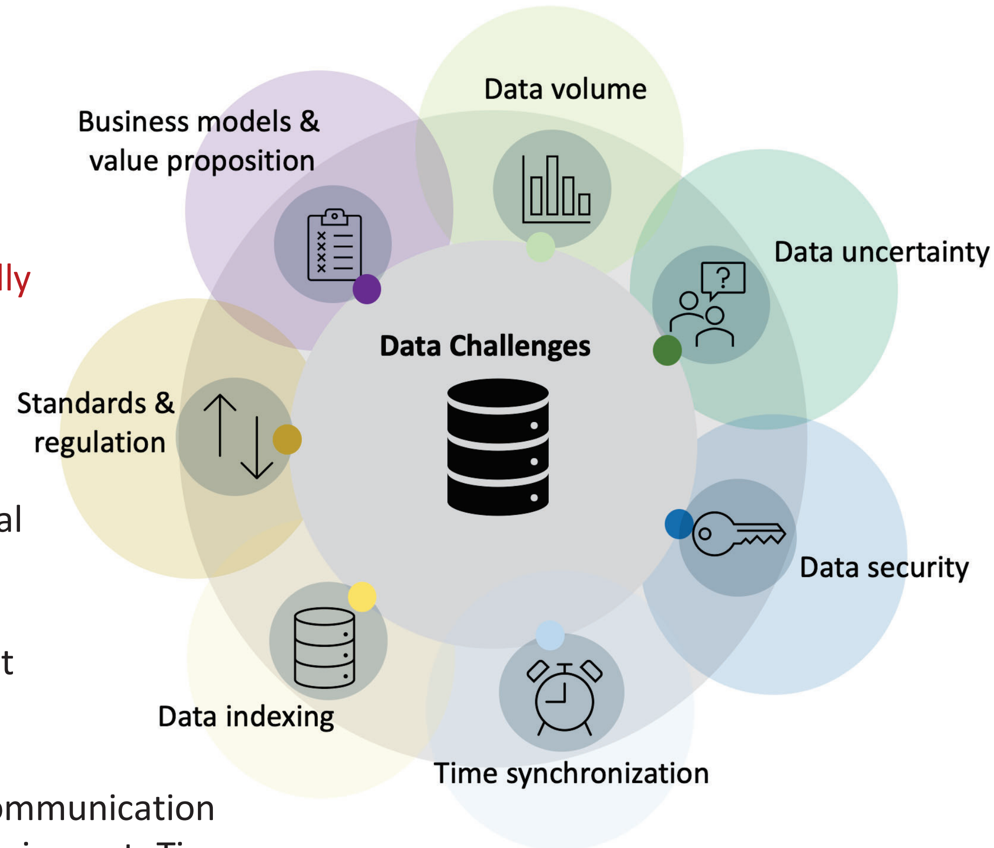
Data Challenges

Volume

- Utilities Data: Grow → Exponentially
- Data Storage, Data Mining, Data processing, Data Querying, and Data Indexing: Grow → Unprecedentedly

Uncertainties

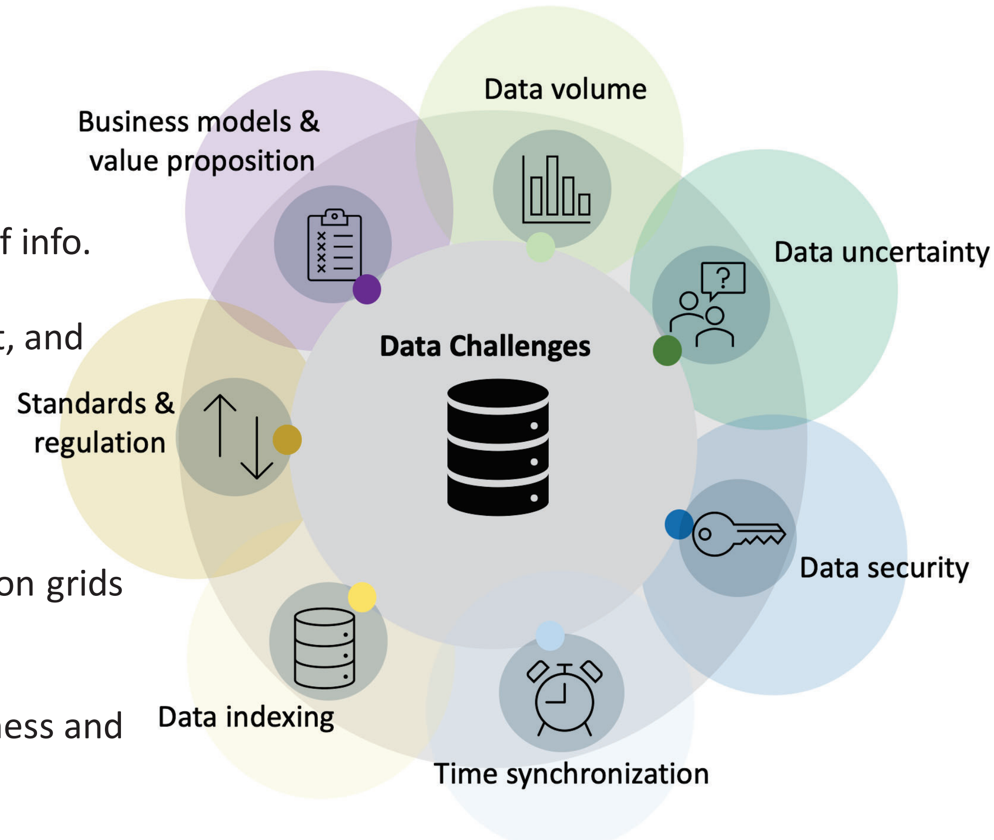
- Characteristics → Real-world smart grid data: lack of data or an incomplete understanding of the operational context
- Operation → Susceptible: Noises, Missing, Inconsistent data → No 100% Certainty
- Major causes: Sensor inaccuracies and imprecision, Communication latencies/delays, Cyber-attack, Physical damages of equipment, Time unsynchronized data, Missing/inconsistent data, noises etc.



Data Challenges

Security

- Data Privacy → Power consumption → Behavior.
- Data Integrity → Prevent unauthorized modification of info.
- Data Authentication → Encryption, trust management, and intrusion detection.



Synchronization

- Real-time control and communication in the distribution grids
- Meaningful connections between events and aids → Forensic analysis of past events, Situational awareness and predictive decisions.
- Unsynchronized Right Now: Communication, storage, and analysis of streams of data from most of the distribution system devices and customers are currently.

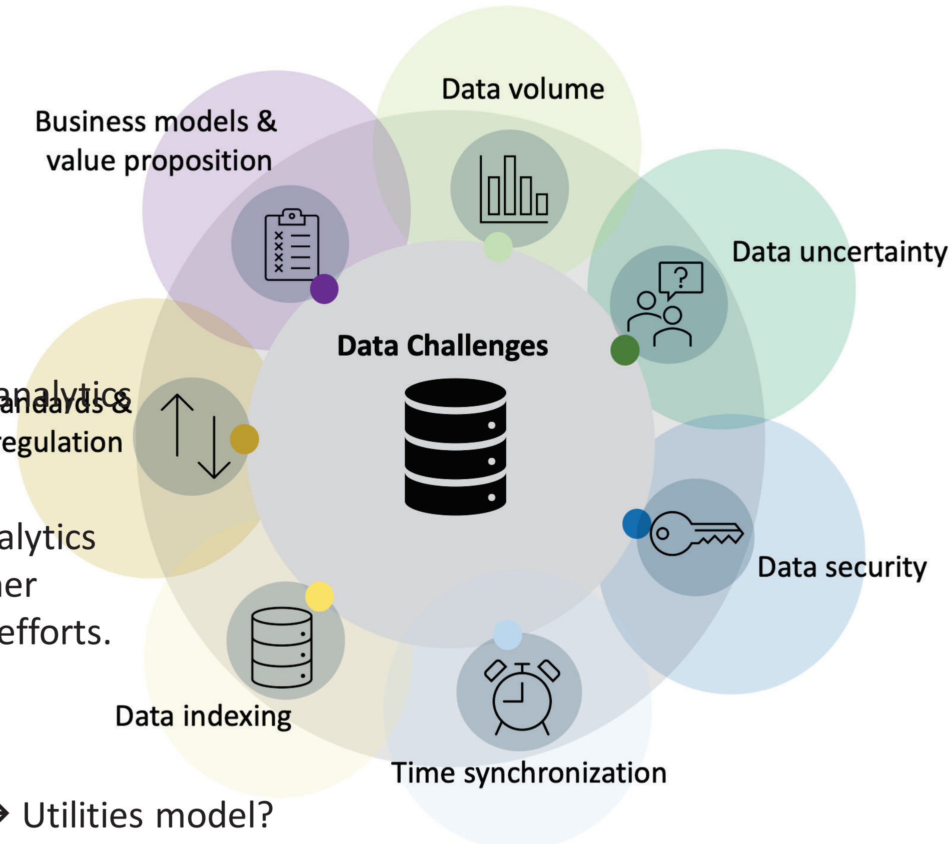
Data Challenges

Data Indexing

- Generic tools (SQL server) and SAP for query purposes
  → Insuffice from the distribution grids application, e.g., real-time applications from big data analytics.
- Advanced data indexing + Query-processing algorithms.

Standards and Regulation

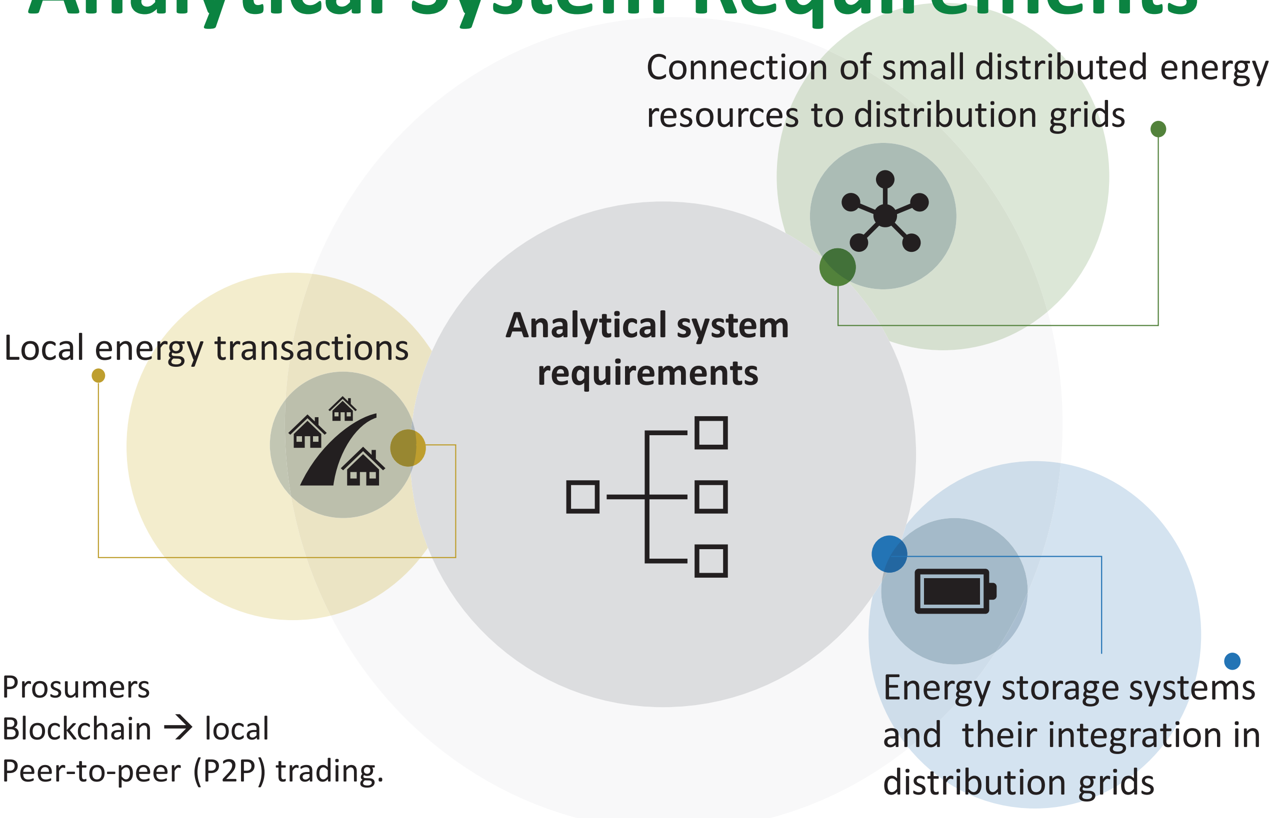
- Different utilities → Different implementing → Big data analytics + Different storage, computing, processing platforms.
- Diversified protocols, architectures, and platforms
  → Limit its potential + Delay the adoption of big data analytics
- Utilities → Do not share data/information with each other
  → Regulatory framework → Facilitate sharing and unify efforts.



Business Models and Value Proposition

- Other industries (e.g., Google, Facebook, and Amazon) → Utilities model?
- How to cash the value?
- Long-term economic and technical values of big data? All stakeholders: Utilities, system operators, and customers.

Analytical System Requirements



Electric mobility
Microgeneration systems
Energy storage systems
Vehicle-to-Grid (V2G) mode

Prosumers
Blockchain → local
Peer-to-peer (P2P) trading.

P2P trading to buy and sell energy services
Provide services to microgrids within utility grids
Supply services to an isolated microgrid.

Diversity of Technologies
Voltage levels
Uncertainties
Whenever possible,
Installed near renewable energy sources
Near end users

Introduction to Data Challenges

Analytical System Requirements

Examples and Solutions

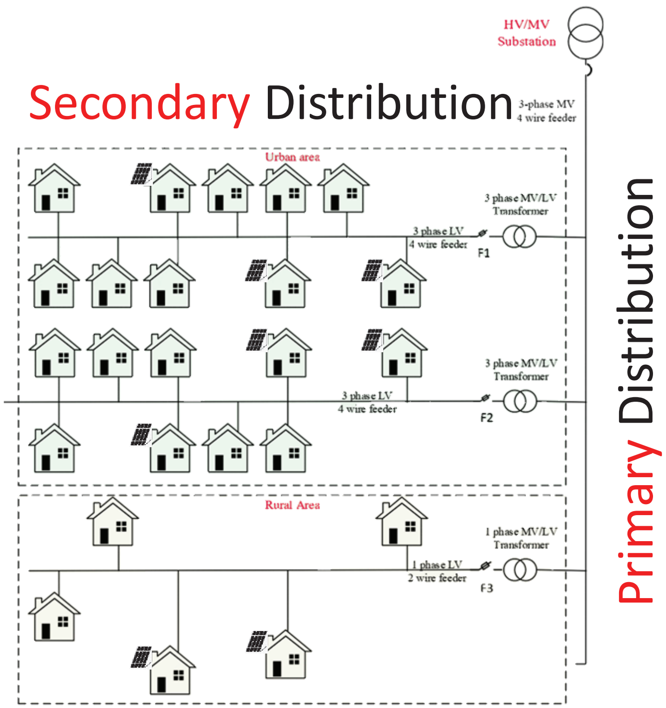
Conclusion

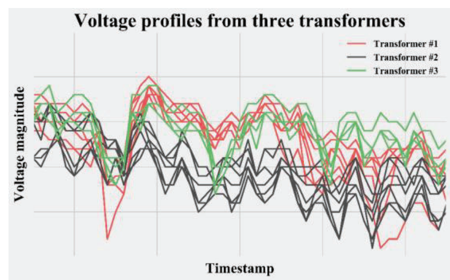
Reference：

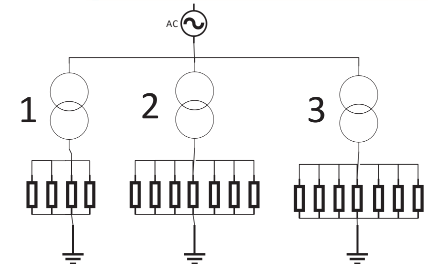
Bhattarai, Bishnu P., et al. "Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions." IET Smart Grid 2.2 (2019): 141-154.

Lopes, João Abel Peças, et al. "The future of power systems: Challenges, trends, and upcoming paradigms." Wiley Interdisciplinary Reviews: Energy and Environment 9.3 (2020): e368.

Lack of Topology & Voltage Data Issues







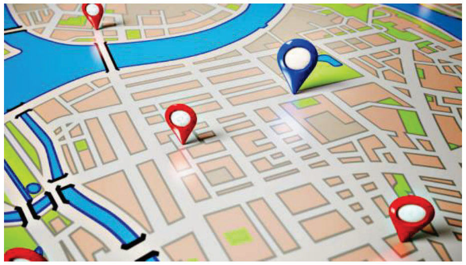
- Sum of downstream meter loads → transformer load
- Transmission-like topology identifier → not possible
  - Distribution systems lack a variety of data measurements

- Voltage variations are a local phenomenon
- However, distant meters may have similar voltages due to similar neighborhood consumption
- Therefore, voltage data alone is not sufficient.

Bad Voltage Data and Data Resolution Inconsistency

- Difference in the voltage and power datasets' temporal resolution.
  - Due to bandwidth limitation
- In addition, the measurements between different meters may not be time synchronized.
- At least 2% of the datasets have missing values or unrealistic values.
  - Bad data and erroneous measurements.
- Missing entire voltage time-series

Location Data Issues



- Missing Addresses
- Missing Block numbers from Addresses
- Incorrect House Addresses
- Moreover, instruments, e.g., poles, transformers, surge arrestors, capacitor banks, etc., not found at the location mentioned

GIS and Other Data Issues

- GPS within the smart meters is not accurate.
- For example, for some places, the location may differ by few streets.
- Therefore, geocoding is needed.

- Data from utilities may lack explanations
  - datasets may not have column names
  - column names containing abbreviations without any explanations
- Transformers and poles may not have measurements
- Meters send mean values instead of instantaneous values
  - Mean voltage instead of instantaneous voltage
  - Hourly energy instead of power measurements

Consequences for Challenges

Transformer Overload

- Transformers may not have measurements
- Distribution systems lack topology information → transformer loading information unavailable
- Increase in load and PV → increase in forward and reverse power → overloads service transformers

Service Voltage Limit Violations

- Increase in PV → two-way power → voltage limit violation
- Lack of topology → prevents identification of transformers and capacitor banks.
- Therefore, voltage violations cannot be corrected.

Power System Analysis

Hosting Capacity Analysis

- Due to challenges, some important power system analyses are impossible
- Hosting Capacity Analysis computes voltages, currents, and power dissipated for each node → finds the maximum amounts of solar PV that can be accommodated.
- However, it needs the topology of distribution system.

State Estimation

- State Estimation is hard in distribution systems, especially the secondary
  - Using pseudo measurements to cover the lack of sensors
- However, similar to hosting capacity analysis,
  - State Estimation is difficult where topology information is unavailable.

Solutions

No Measurements at Transformers or Poles

- Usually, transformers and poles do not have any voltage measurements.
- Therefore, one can obtain meter clusters supplied by a common transformer.
- Next, meter clusters are mapped to the transformers using GIS information.
- Such an approach is more accurate than nearest neighbors.

Lack of Topology Info

Meter-Transformer Mapping

- If topology information is not available, we recover the meter-transformer mapping.
- Sum of downstream meter loads → load on parent transformer.
- Can shift the load from overloaded transformers → less-loaded transformers.

Lack of Topology Information

PV-Feeder Mapping

- Similarly, we recover the PV-feeder mapping.
- Identify the feeder with potential reverse power.
- Adjust the transformer tap settings and/or the capacitor banks to fix voltage.

Voltage Similarity Issues

Use of Location Data

- Location data + voltage data → resolve distant voltage similarity
- For example, GPS information can be considered with the Haversine metric for computing distances.
- Distant meters with similar voltage
  - Voltage Euclidean distance → 0
  - Haversine distance ≠ 0

Meter Location Inaccuracy Issue

Geocoding Addresses

- Each meter has a house address which is available to the utilities.
- Instead of using GPS of meter, one can geocode smart meter address to obtain accurate location
- Online Geocoding has a fee, but it is a one-time cost for the utility.

Other Location Data Issue

- To obtain accurate locations for poles and transformers, one can identify their GPS coordinates using Google Street-View.
- Such locational data has high accuracy
- The service is free and it does not require any field visit

Missing Data Values and Unrealistic Values

- Interpolating for the missing values → Introduces noise
- Removing such timesteps may leave limited data
- A better idea is to process the data in batches and ignore the timesteps involving missing values or unrealistic values

Introduction to Data Challenges

Analytical System Requirements

Examples and Solutions

Conclusion

Reference：

Bhattarai, Bishnu P., et al. "Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions." IET Smart Grid 2.2 (2019): 141-154.

Lopes, João Abel Peças, et al. "The future of power systems: Challenges, trends, and upcoming paradigms." Wiley Interdisciplinary Reviews: Energy and Environment 9.3 (2020): e368.

Conclusions

- Identification of the meter-transformer mapping → transformer loading information → shift load from overloaded to less loaded transformers.
- Identification of the PV-feeder mapping → feeder power generated → Adjust the transformer tap settings and/or the capacitor banks to fix voltage.
- Location data + voltage data → resolve distant voltage similarity
- Geocoding can resolve meter location inaccuracy issue easily
- Google Street-View → Accurate poles and transformer locations
- Process the data in batches and ignore the timesteps involving missing values or unrealistic values