# MODULE 2: Role of Big Data in Smart Grid Communications:

## Structure

2.1 Introduction,

2.2 The Grid Modernization,

2.3 The Grid Interconnection with the Internet of Things,

2.4 Data Traffic Pattern in a Smart Grid Environment,

2.5 The Massive Flow of Information in a Smart Scenario,

2.6 The Volume of Generated Data in a Smart Distribution System: A Case of Study.

## Objectives of Course

### Learning Objectives

To explain the role of big data in smart grid communications and optimization of big data in electric power systems.

## **2.1** Introduction

- Smart Grid as a Sensor Network – Consists of many connected smart devices.
- Rising Data Flow – Growth of devices and operational needs increases information exchange.
- Big Data Role – Provides efficient management of large data volumes.
- Benefits – Helps utilities understand customer behavior, demand, weather, outages, and failures.
- Robust Methodologies – Needed to make the grid smarter and more reliable.
- Key Concern – Quantifying and utilizing data from devices effectively.
- Chapter Aim – To characterize and evaluate data growth in smart grid communication networks.
- **Example – Active distribution system shows massive data generation for monitoring and control.**

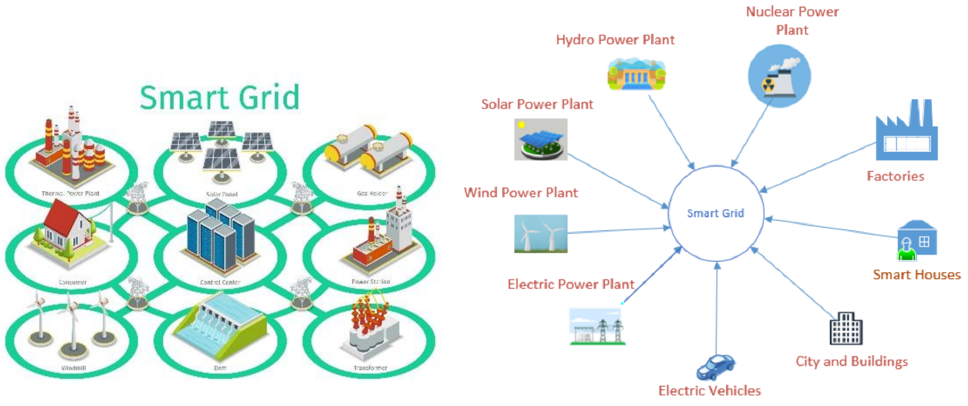



**Fig-1: Smart Grid**

## 2.2 The Grid Modernization,

- **Dynamic Nature –** Power generation variations directly affect customer supply → requires real-time monitoring and operation.
- **Smart Grid –** Integrates ICT with electrical system to improve reliability, robustness, safety, reduce losses and failures (especially at distribution level where 90% failures occur).
- **Distributed Generation**

  **(Solar/Wind) –** Higher penetration increases need for monitoring and efficient info exchange among all agents.

- **Role of IEDs –** Real-time automation through data from user profiles, PMUs, distributed resources, AMI, sensors, actuators, breakers, capacitor banks, and utilities.
- **Advanced Metering Infrastructure (AMI) –** First step in modernization; enables two-way communication between users and utilities, supports demand response, pricing, state estimation, and real-time supply-demand balance.
- **Phasor Measurement Units (PMUs) –** Applied in protection, state estimation, voltage/frequency control, renewable integration, and islanding monitoring.
- **Massive Data Generation** – Intelligent devices and sensors create huge volumes of information requiring proper management.
- **Big Data as Solution** – Provides tools to handle and analyze large datasets, improving decision-making, grid robustness, and profitability.
- **Key Requirement** – Quantify which devices generate data and how much, to effectively apply Big Data technologies.

## 2.3 The Grid Interconnection with the Internet of Things

- **Intelligent Communication Network** – Must be high-performance, reliable, robust, flexible → **handles data collection, routing, monitoring, and management.**
- **Scale of Communication** – Thousands of devices send data to substations and control centers with minimal human interaction; each distribution system has **hundreds–thousands of smart meters, PMUs, and IEDs.**
- **IoT in Smart Grids** – Enhances flexibility, scalability, and automation → forms a cyber-physical system integrating power and communication networks.

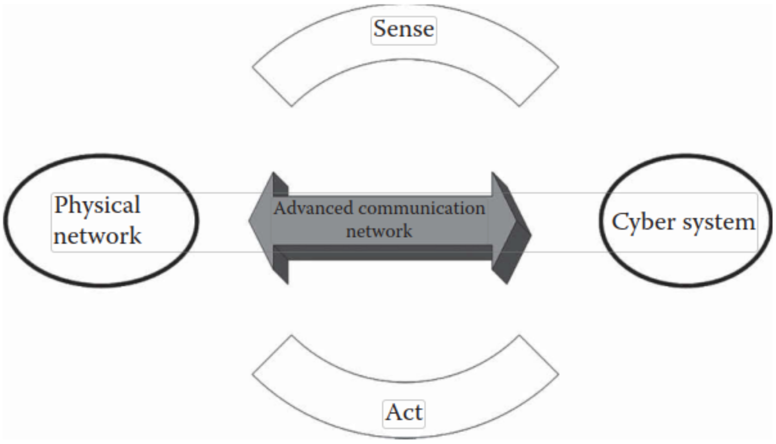



**Fig:**Smart grid as a cyber–physical system. Adapted from Yu and Xue (2016).

- **Key Communication Requirements**:
- **Latency** – network delay must be minimal.
- **Bandwidth** – critical for choosing wired/wireless tech.
- **Interoperability/Flexibility** – systems must work together.
- **Data Throughput** – max info transfer capacity.
- **Cybersecurity** – authentication, authorization, privacy.

- **Data Characteristics** – IoT devices generate big data: heterogeneous, varied, unstructured, noisy, redundant.

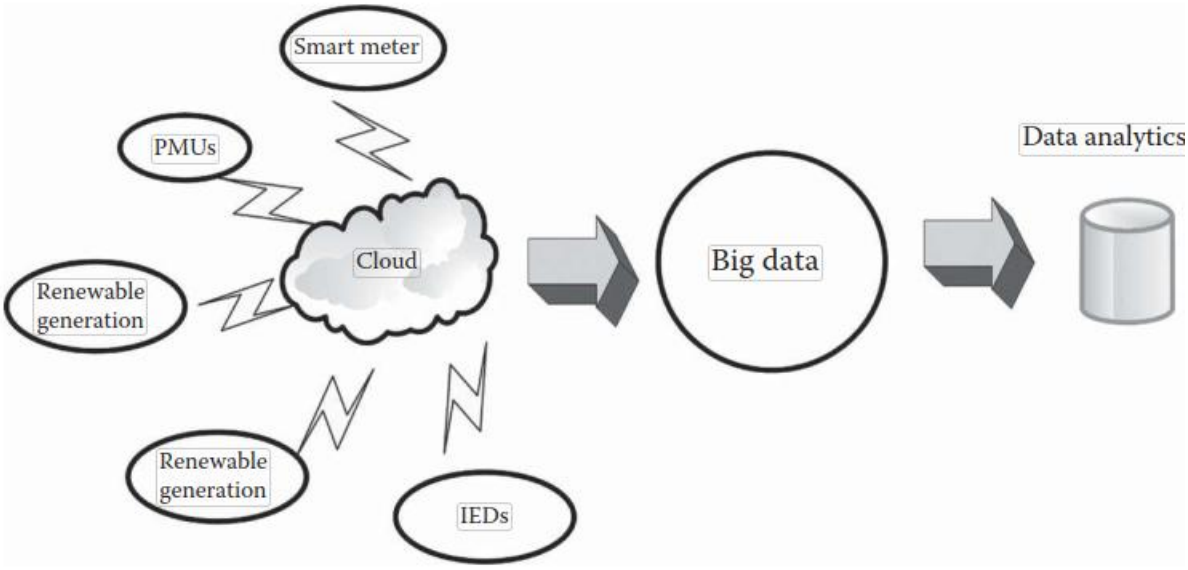



**Fig** :The integration of IoT and Big Data technology. Adapted from Marjani et al. (2017).

- **Technology Choices** – Must be reliable, secure, cost-effective, and adaptable to future needs.
- **Wired** – high data rate, costly, low mobility.
- **Wireless/Mobile** – scalable, flexible, lower cost, supports **M2M communication**.
- **Mobile Standards** – LTE-based enhancements for IoT.
- **NB-IoT (3GPP Release 13 & 14)** – efficient spectrum use, wide coverage, low-cost devices, high capacity, supports massive machine-to-machine (M2M) communication.

## 2.4 Data Traffic Pattern in a Smart Grid Environment

- Traditional power systems are **unidirectional**, with limited intelligence and no real-time communication (Bouhafs et al., 2012).
- Smart grids require **advanced communication systems** for protection, monitoring, and control.
- Existing systems like **SCADA** and **AMR** are inadequate for future smart grid demands (Lai and Lai, 2015).
- Future grids need **bidirectional communication**, ensuring robustness, reliability, and security.

**Main functions** (Marjani et al., 2017; Fan et al., 2012):

1. **Distribution Control & Protection** – **IED devices detect/locate faults** and exchange reporting messages.
2. **Wide Area Monitoring System** – Collects info from substations and large areas for critical decisions.
3. **Demand Response (DR)** – Manages multiple distributed energy resources with intermittent variables.

4. **Advanced Metering Infrastructure (AMI)** – Smart meters for billing, consumer interaction, load control, DR, and islanding detection.

- Each application has **different latency and data requirements** (Kuzlu et al., 2014):
- **Protection**: 1–10 ms, few bytes
- **Control**: ~100 ms, few bytes
- **Monitoring**: ~1 s, few–medium bytes
- **Metering**: minutes–hours, medium size
- AMI enables functions like **choosing energy sources, demand monitoring, and dynamic billing** (Sánchez-Ayala et al., 2013).
- Smart grids will produce **huge amounts of data**, e.g., PMUs (sampled at 10–60 per cycle) and AMI (every 1–15 min).
- **Big data analytics** is crucial for maintaining reliability, robustness, and grid modernization (Daki et al., 2017).

**Phasor Measurement Units (PMUs)**

- PMUs placed strategically perform **accurate voltage and current phasor measurements** using GPS (Liu et al., 2012).
- Data is sent to **Phasor Data Units (PDUs)** for processing.
- PMUs generate measurements per cycle (10, 20, 30, 60 most common).
- Required bandwidth depends on:
- **BW = Nframe × fs × NPMU**
    - Nframe = frame size (bytes)
    - fs = sampling frequency
    - NPMU = number of PMUs connected
- **Higher sampling frequency** = **higher transmission capacity needs**.
- **Increasing PMU deployment** = **more synchrophasor data to analyze**.

**Advanced Metering Infrastructure (AMI)**

- Smart meters enable **numerous smart applications** when paired with reliable communication infrastructure.
- Each meter generates data with **different message size and sample rates** (Luan et al., 2010).
- Smart meters will be deployed in **residential, commercial, and industrial facilities**, creating massive data flow.
- **Estimating number of smart meters** depends on communication network design (coverage, antenna gain, propagation losses, geography).
- Formula for number of smart meters:

**NSM = $\rho \pi d^2$**

    - $\rho$ = smart meter density (per $m^2$)
    - d = coverage radius of base station (urban, suburban, rural variations).
- **Minimum data rate per smart meter = 64 kbps** for critical operations (Persia et al., 2015).

## 2.5 The Massive Flow of Information in a Smart Scenario

- **Massive data generation**:
  - Growth of **IEDs** (Intelligent Electronic Devices) → exponential data flow.
  - Utilities already generate **hundreds of millions of gigabytes**, rising to **terabytes annually**.
  - **Synchrophasors alone** produce **hundreds of terabytes per year** (Asad et al., 2017).
- **Opportunities from big data**:
  - Improves **grid reliability** and enables applications such as:
    - **Predictive analytics**
    - **Demand-side management**
    - **Real-time grid awareness**
    - **Outage detection**
    - **Asset management**
    - **Theft detection**

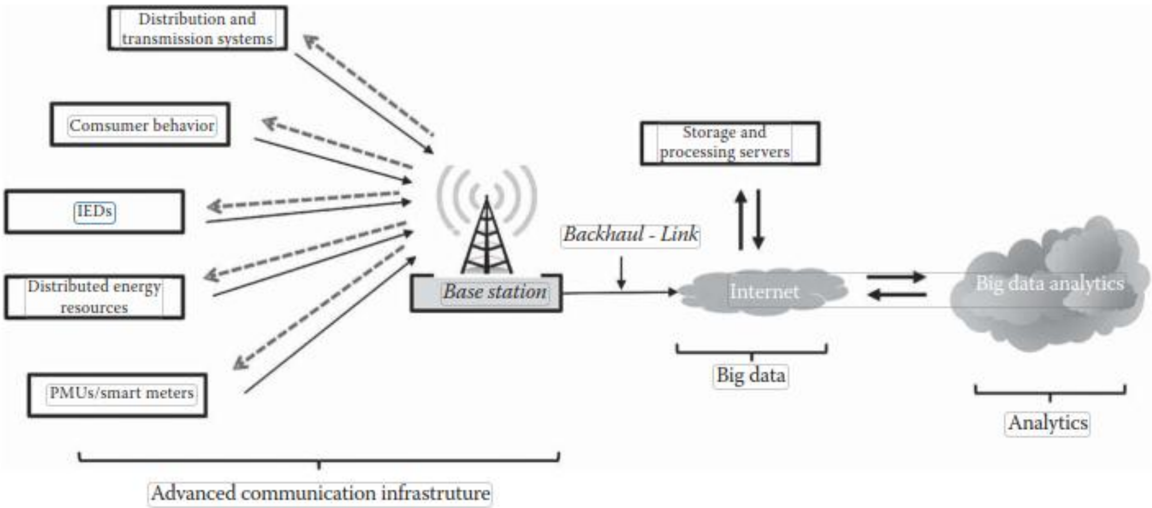



**FIGURE 3.3**
The integration of information and communication technologies in a smart grid context.

**Role of big data** (Jiang et al., 2016):

- Enhances **prediction, management, and processing** of grid operations.
- Recognizes **patterns in data** for better operational decisions.
- <u>**Key application areas:**</u>
  - **Demand Response (DR):** Predicting/analyzing user patterns → accurate demand forecasting.
  - **Distributed Energy Resources (DER):** Forecasting & scheduling intermittent energy sources for planning.
  - **AMI:** Smart meter data → customer behavior, load forecast, demand management.
  - **Distribution Automation:** Monitoring and sensing → predict outages, improve robustness.

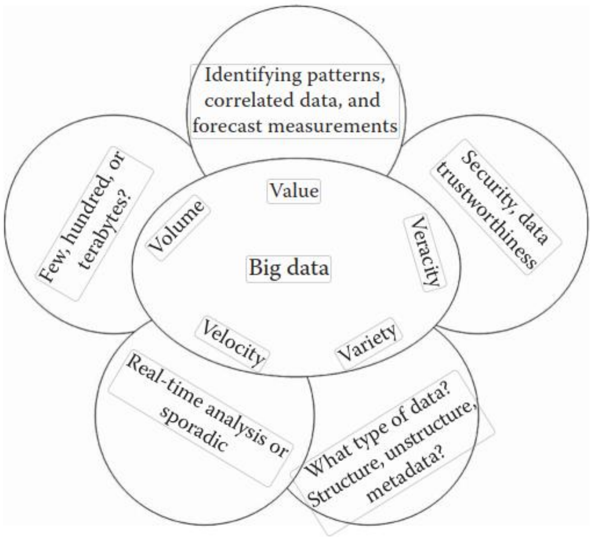



**FIGURE 3.4**
The 5 V's of big data in a smart grid scenario. Adapted from Subhani et al. (2015).

**Big data foundation – 5 V's** (Subhani et al., 2015):

- **Volume** – massive data scale
- **Velocity** – high-speed generation & transfer
- **Variety** – diverse data formats (structured/unstructured)
- **Value** – usefulness of data for decisions
- **Veracity** – data quality and accuracy

**Advanced data analytics methods**:

- Predictive analytics, data mining, AI, fuzzy theory

## 2.6 Volume of Generated Data in Smart Distribution – Case Study

**General**

- **Future smart grids = huge data** from IEDs, PMUs, and smart meters.
- In normal mode: periodic traffic (e.g., metering every 15 min).
- In critical mode: continuous messages, DR, load flow, outage info.

**Case Study I – Data Generated by PMUs**

**Step-1: Setup**

- IEEE 123 bus system is selected (high-load topology, widely used in research).
- PMU placement follows Jamil et al. (2014).
- **49 PMUs** deployed → optimal number for this system.

**Step-2: Data Frame Structure**

- Each PMU sends **frame packets** as per IEEE Std. C37.118 (2011).
- Packet = **80 fixed bytes** + fields (phasors, transducers, digital signals).
- Each PMU has **8 phasor channels** + **2 digital channels**.

**Step-3: Sampling Rates**

- Brazilian system frequency = **60 Hz**.
- Sampling rates considered: **10, 20, 30, 60 synchrophasors/sec**.

- Higher sampling rate = higher data generation.

**Step-4: Bandwidth Requirement Formula**

$$BW = N_{frame} \times f_s \times N_{PMU}$$

- Nframe = frame size (bytes)
- fs = sampling frequency
- NPMU = number of PMUs (49 in this case)

**Step-5: Results**

- At 10 samples/sec → data volume already in **hundreds of Mbps**.
- At higher rates → data grows drastically.

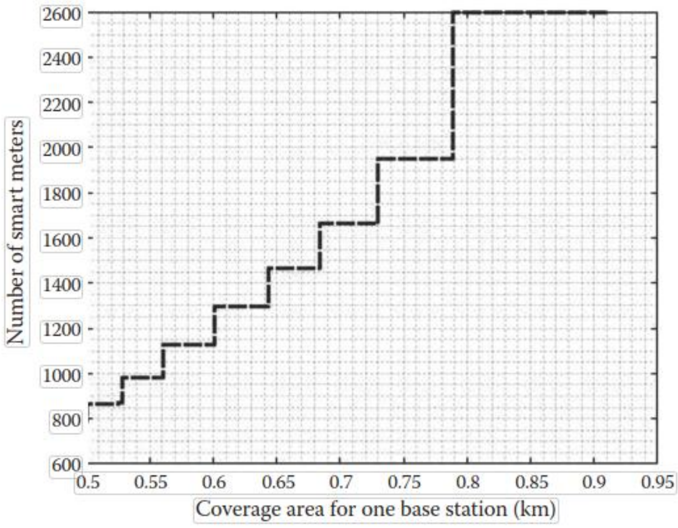
**Step-6: Inference**

- PMUs generate **massive continuous data streams**.
- Without **big data analytics**, it is difficult to handle, store, and process.
- Organized analysis = better grid performance and reliability.

Simulation: **IEEE 123 bus system** with NB-IoT communication.

Volume of Generated Data by PMUs

| | | Sampling Rate (Mbps) | | | | | |
|---|---|---|---|---|---|---|---|
| **#PMU** | **#Bytes** | **10** | **12** | **15** | **20** | **30** | **60** |
| 1 | 112 | 8.96 | 10.752 | 13.44 | 17.92 | 26.88 | 53.76 |
| 2 | 224 | 16.06 | 19.27 | 24.08 | 32.11 | 48.17 | 96.34 |
| 3 | 336 | 43.16 | 51.79 | 64.74 | 86.32 | 129.48 | 258.96 |
| 6 | 672 | 232.02 | 278.43 | 348.04 | 464.05 | 696.07 | 1392.15 |
| 9 | 1008 | 1871.05 | 2245.26 | 2806.57 | 3742.10 | 5613.15 | 11226.29 |
| 10 | 1120 | 16764.60 | 20117.51 | 25146.89 | 33529.19 | 50293.79 | 100587.57 |
| 15 | 1680 | 225316.16 | 270379.39 | 337974.24 | 450632.32 | 675948.47 | 1351896.95 |
| 20 | 2240 | 4037665.548 | 4845198.66 | 6056498.323 | 8075331.097 | 12112996.6 | 24225993.29 |
| 30 | 3360 | 108532449.9 | 130238940 | 162798674.9 | 217064899.9 | 325597350 | 651194699.7 |
| 49 | 5488 | 4765008682 | 5718010419 | 7147513023 | 9530017364 | 1,4295E+10 | 28590052093 |
| 50 | 5600 | 2,13472E+11 | 2,5617E+11 | 3,20209E+11 | 4,26945E+11 | 6,4042E+11 | 1,28083E+12 |
| 60 | 6720 | 1,14763E+13 | 1,3772E+13 | 1,72144E+13 | 2,29526E+13 | 3,4429E+13 | 6,88577E+13 |




**Case Study II – Data Generated by AMI (Smart Meters)**

**Step-1: Setup**

- Smart meters deployed across the same **IEEE 123 bus system area (≈2 km², urban scenario)**.
- Meters randomly placed, exact optimal allocation not studied here.

**Step-2: Data Rate Assumption**

- Each smart meter requires **64 kbps** in critical mission scenarios (Persia et al., 2015).

**Step-3: Coverage Analysis**

- Communication via **NB-IoT wireless**.
- Factors: distance, propagation loss, antenna gain, coverage area.
- One base station coverage = **0.5 km²** → supports ~**800 meters**.

**Step-4: Scaling to Whole System**

- Total area = ~2 km².
- Need **5 base stations** to cover entire system.
- Supports ~**4000 smart meters**.

**Step-5: Inference**

- More smart meters = **exponential data growth**.
- Data must be processed by **advanced analytics** to avoid overload.
- Correct treatment ensures **efficient grid + communication flow**.

**Inference**

**Case I (PMUs):** Generate <u>**huge continuous data**</u>; needs big data analytics for reliability.

**Case II (AMI):** Thousands of smart meters → **exponential data; requires efficient communication & analytics.**

- **Overall:** Smart grids produce massive data; without proper treatment it's a burden, with analytics it ensures reliability and efficiency.

## Outcomes

At the end of the module, students will be able to:
CO-1: Interpret the role of big data and machine-learning methods applicable to power systems and in particular to Smart Grid communications. [L2]

**TEXT BOOKS:**

**Big Data Analytics in Future Power Systems, Ahmed F. Zobaa and Trevor J. Bihl, CRC Press 2019. 2019.**

**Reference Books/ Link**

1. **Big Data Analytics for Power Systems – Big Data Analytics in Power Systems**
2. **Application of Big-Data Analytics in Power System Protection-Lec-37: Application of Big-Data Analytics in Power System Protection**