The Big Data Analytics: An Incremental Improvement or a Transformational Change

Mladen Kezunovic, Ph.D., P.E.
Regents Professor
Texas A&M University

July 29, 2020

Outline

Example: Outage Management
What is Incremental and what is Transformational?
Transformational Aspects:
- Data Sources and Data Management
- Data Analytics (machine learning)
- Risk Prediction
- Risk Management and Mitigation
- Benefits
Future Decision-Making Framework
Other Examples
Conclusions
How Power Outages Occur?

- Lightning damaging insulators
- Trees touching lines/feeders
- Equipment wear and tear
- Animal intrusions
- Infrastructure destroyed due to natural disasters
- Human-caused reasons
- Failure of fault clearing equipment
- Insulation stress going over limits
Why Outages are Becoming More Frequent

Observed Outages to the Bulk Electric System, 1992-2012

- Weather-Related
- Non-Weather Related
- Unclassified

Source: Energy Information Administration

FIGURE 1. U.S. Electric Grid Disruptions

- All Events
- Weather-related Events
- All Events Trendline
- Weather-related Events Trendline

Electric Disturbance Event Date

The Department of Energy tracks major electric disturbance events through Form GE-417. Utilities submit information about qualifying incidents, including when they occurred, where they occurred, what triggered them, and how many customers were affected. Notably, while the reported number of non-weather-related events is high, the vast majority of incidents resulting in customer outages occur because of weather.

©2020 Mladen Kezunovic, All Rights Reserved
Outages Affected by Severe Weather

Severe Weather Caused 80% of Large-Scale Power Outages and 88% of Customers Affected by Power Outages From 2003-2012

Figure 3. Percentage of power outages and total customers affected caused by severe weather from 2003-2012.

General Storms Caused Nearly 60% of All Weather-Related Outages Between 2003-2012

Figure 6. Types of severe weather responsible for all weather-related power outages from 2003-2012.

- 59% Storms & Severe Weather (including heavy rain, thunderstorms, high winds, lightning)
- 18% Cold Weather & Ice Storms
- 18% Hurricanes and Tropical Storms
- 3% Tornadoes
- 2% Extreme Heat & Wildfires
Outage Causes

What causes our power outages?

- 2015
  - Mother Nature: 35.89%
  - Equipment: 18.3%
  - Vegetation: 16.2%
  - Other: 15.5%
  - Human: 7.7%
  - Animals: 6.3%
- 5-yr. average
  - Mother Nature: 27.8
  - Equipment: 25.4%
  - Vegetation: 23.9%
  - Other: 15.8%
  - Human: 7.7%
  - Animals: 6.1%

Source: Alaska Electric light and Power Company

- Weather: 27%
- Fallen trees: 20%
- Equipment: 29%
- Miscellaneous: 7%
- Human: 7%
- Animal: 11%

Source: We Energies

Major causes of power outages in the U.S.

- Weather/Tree-related: 82%
- Equipment failure: 15%
- Unknown/Other: 10%
- Public or Animal contact: 7%
- Power Grid failure: 5%
- Maintenance: 1%

Source: Annual Eaton Investigation 2013
Outline

Example: Outage Management

What is Incremental and what is Transformational?

Transformational Aspects:
- Data Sources and Data Management
- Data Analytics (machine learning)
- Risk Prediction
- Risk Management and Mitigation
- Benefits

Future Decision-Making Framework

Other Examples

Conclusions
What is Incremental and what is Transformational?

<table>
<thead>
<tr>
<th>Property/method</th>
<th>Incremental</th>
<th>Transformational</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Sources</td>
<td>Limited to the grid measurements and models</td>
<td>Much more broader data from multiple sources within and outside utilities</td>
</tr>
<tr>
<td>Data Management</td>
<td>Use of legacy platforms for field data collection (Digital relays, DFRs, PMUs)</td>
<td>New data platforms allowing data ingestion, curation, cleansing, wrangling, etc.</td>
</tr>
<tr>
<td>Data Analytics</td>
<td>Physics-based</td>
<td>Data-based</td>
</tr>
<tr>
<td>Results</td>
<td>Corrective</td>
<td>Predictive</td>
</tr>
<tr>
<td>Mitigation</td>
<td>Reactive (after the fact)</td>
<td>Preventive</td>
</tr>
</tbody>
</table>
Outline

Example: Outage Management

What is Incremental and what is Transformational?

Transformational Aspects:
- Data Sources and Data Management
- Data Analytics (machine learning)
- Risk Prediction
- Risk Management and Mitigation
- Benefits

Future Decision-Making Framework

Other Examples

Conclusions
Data Sources

- Utility measurements
- Animals Data
- Weather Forecast
- Network Assets Data
- Lightning Data
- Vegetation Indices
- Market data
- GIS
- UAS

©2020 Mladen Kezunovic, All Rights Reserved
Vegetation Data

Environmental Impacts:
- precipitation
- wind
- temperature
- humidity
- lightning

WIND

GROWTH

- Tree Trimming Performed
- Outage
- Reactive Tree Trimming
- Periodic Tree Trimming Scheduled

TIME

WEATHER MEASUREMENTS
- temperature
- wind speed
- gust and direction
- precipitation
- humidity
- pressure
- lightning parameters

VEGETATION MAP FROM HIGH RESOLUTION IMAGERY
- distance to the lines
- growth rate
- canopy height
- canopy spread
- health index
- tree species
Weather Data

Weather Station

Radar

Satellite

Example: Apparent Temperature
Data download: every 3 hours
Forecast for next 3 days
Data resolution: 3 hours

National Digital Forecast Database (NDFD)
Data Management

Figure 1. Vegetation parameters extraction

Table I. Vegetation data sources and datasets

<table>
<thead>
<tr>
<th>Source</th>
<th>Data set name</th>
<th>Spatial res.</th>
<th>Contains</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Oceanic and Atmospheric Administration</td>
<td>2010 C-CAP Regional Forest Fragmentation Land Cover, [12]</td>
<td>30 m</td>
<td>Forest fragmentation</td>
</tr>
<tr>
<td>Florida Division of Emergency Management</td>
<td>2007 St. Lucie LIDAR, [13]</td>
<td>Vertical 0.6 ft Horizontal 3.8 ft</td>
<td>LIDAR data</td>
</tr>
<tr>
<td>National Aeronautics and Space Administration</td>
<td>MODIS - Moderate Resolution Imaging Spectroradiometer, [14]</td>
<td>0.5 km</td>
<td>Vegetation indices</td>
</tr>
<tr>
<td>National Cooperative Soil Survey</td>
<td>gSSURGO, [15]</td>
<td>10 m</td>
<td>Soil data</td>
</tr>
</tbody>
</table>
Data Analytics: Risk Formulation

\[ R(X,t) = P[T(X,t)] \cdot P[C(X,t)|T(X,t)] \cdot u(C(X,t)) \]

Intensity \( T \) – Weather severity

Hazard – Probability of a severe weather impact

Vulnerability – Probability of a vegetation caused outage for a given hazard

Economic Impact – Cost of periodic and reactive tree trimming

Hazard and Vulnerability

Table II. Impact of wind on vegetation

<table>
<thead>
<tr>
<th>WS [m/s]</th>
<th>Effect on trees</th>
<th>Hazard</th>
<th>WS [m/s]</th>
<th>Effect on trees</th>
<th>Hazard</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.5-7.9</td>
<td>Small branches movement</td>
<td>0</td>
<td>17.2-20.7</td>
<td>Twigs broke from trees</td>
<td>2</td>
</tr>
<tr>
<td>8.0-10.7</td>
<td>Moderate size branches move</td>
<td>0</td>
<td>20.8-24.4</td>
<td>Large branches broke from trees</td>
<td>3</td>
</tr>
<tr>
<td>10.8-13.8</td>
<td>Large sized branches</td>
<td>1</td>
<td>24.5-28.4</td>
<td>Trees uprooted or broken</td>
<td>4</td>
</tr>
<tr>
<td>13.9-17.1</td>
<td>Whole trees in motion</td>
<td>1</td>
<td>&gt;28.5</td>
<td>Severe vegetation damage</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 2. Land-based weather stations – processing of historical weather data

Figure 3. Weather forecast - data example and processing
Risk Prediction

Weather Hazard

Network Vulnerability

Risk Map

Overall risk reduction 32.85%
Reactive tree trimming cost reduction 27.2%
Risk Management and Mitigation: Optimal Tree Trimming

For total of N feeder sections maximize the reduction in vegetation risk

$$\max \left\{ R = \sum_{t=1}^{T} \frac{1}{N} \sum_{n=1}^{N} \Delta R_{n,t} \cdot F_{n,t} \right\}$$

$$F_{n,t} = \begin{cases} 0, & \text{section not trimmed} \\ 1, & \text{section trimmed} \end{cases}$$

Constrains:
- Total cost of tree trimming limit:
  $$\sum_{t=1}^{T} \sum_{n=1}^{N} F_{n,t} \cdot PC_{n,t} \leq PA$$
- One section trimmed at the time:
  $$\sum_{n=1}^{N} F_{n,t} = 1$$

Difference in component risk before and after action:

$$\Delta R_{n,t} = R_{n,t}^{before} - R_{n,t}^{after}$$

- \(n\) feeder section
- \(N\) total number of feeder sections
- \(t\) time instance
- \(T\) number of time instances
- \(\Delta R_{n,t}\) reduction in risk after trimming
- \(PC_{n,t}\) cost of trimming on one section
- \(PA\) total allocated tree trimming budget
# Risk Management and Mitigation

![Diagram showing environmental impacts: precipitation, wind, temperature, humidity, lightning]

![Vegetation map from high resolution imagery or LiDAR]

<table>
<thead>
<tr>
<th>ID</th>
<th>Zone Order</th>
<th>Average Risk Reduction [%]</th>
<th>Economic Impact Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12,1,21,22,13,24,2,3,10,19,11,5,6,18,4,23</td>
<td>32.18</td>
<td>0.39</td>
</tr>
<tr>
<td>2</td>
<td>12,1,13,24,21,22,2,3,10,19,11,5,6,18,4,23</td>
<td>31.98</td>
<td>0.43</td>
</tr>
<tr>
<td>3</td>
<td>1,12,21,22,10,19,11,5,13,24,2,23,3,6,18,4</td>
<td>26.14</td>
<td>0.28</td>
</tr>
<tr>
<td>4</td>
<td>12,1,24,13,2,3,10,21,11,5,6,18,4,22,19,23</td>
<td>23.84</td>
<td>0.25</td>
</tr>
<tr>
<td>5</td>
<td>1,12,21,22,24,13,3,10,2,19,6,4,11,5,23,18</td>
<td>20.89</td>
<td>0.26</td>
</tr>
</tbody>
</table>
Benefits of Outage Management

Annual Business Losses from Grid Problems

Primen Study: $150B annually for power outages and quality issues

The real victim of power outages are businesses in general
US$ 000 (2010): average cost of one hour power interruption in the US per type of customer

Source: US Department of Energy.
Outline

Example: Outage Management

What is incremental and what is transformational?

Transformational Aspects:
- Data Sources and Data Management
- Data Analytics (machine learning)
- Risk Prediction
- Risk Management and Mitigation
- Benefits

Future Decision-Making Framework

Other Examples

Conclusions
Future Decision-making Framework

- Weather: Temperature, Wind, Precipitation, Clouds, Lightning, Humidity
- Risk Mitigation: Monitoring, Control and Protection, Flexible Load
- Power Generation: Renewables (Hydro, Solar, Wind), Fossils (Gas, Coal)
- Power Transmission: Substations, Lines, Cables
- Power Distribution: Substations, Lines, Cables
- Customers: Residential, Commercial, Industrial

Probabilistic Spatio-Temporal State of Risk

Risk Assessment: Probabilistic Regression, Causal Probabilistic Risk Assessment

Real Time Market
Forward Market
Outline

Example: Outage Management
What is Incremental and what is Transformational?
Transformational Aspects:
  - Data Sources and Data Management
  - Data Analytics (machine learning)
  - Risk Prediction
  - Risk Management and Mitigation
  - Benefits
Future Decision-Making Framework
Other Examples
Conclusions
Transmission Outage Prediction

Historical Weather Data

Fractions of missing data from ASOS observations

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Dew Point</th>
<th>Humidity</th>
<th>Wind Direction</th>
<th>Wind Speed</th>
<th>Precipitation</th>
<th>Pressure</th>
<th>Wind Gust</th>
<th>Weather Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.146</td>
<td>0.148</td>
<td>0.148</td>
<td>0.145</td>
<td>0.134</td>
<td>0.312</td>
<td>0.265</td>
<td>0.378</td>
<td>0.336</td>
</tr>
</tbody>
</table>
Weather Forecast

No Outage - Precipitation

Outage - Precipitation

No Outage – Wind Speed

Outage – Wind Speed
Risk Maps

• **No outages occurred** ⇒ outage probabilities are **smaller than 60%** for all substations

• **Outages occurred** ⇒ the area around the outages has points with probability over 80%
Distribution Transformer Failure Prediction

The System – South Korea

Distribution Facilities in JeonllaNam-do Area

<table>
<thead>
<tr>
<th>Transformer Bank</th>
<th>Number</th>
<th>Capacity (kVA)</th>
<th>Breaker</th>
<th>Equipment</th>
<th>COS</th>
</tr>
</thead>
<tbody>
<tr>
<td>104</td>
<td>243</td>
<td>9,252</td>
<td>12</td>
<td>1.4</td>
<td>72</td>
</tr>
</tbody>
</table>

The comparison of the number of DT in South Korea
# Weather Data

![Historical Weather Measurements](https://www.ncdc.noaa.gov/nexradinv/chooseday.jsp?id=rksg)

<table>
<thead>
<tr>
<th>Lightning [0/1] (LI)</th>
<th>Average Temperature [°F] (AT)</th>
<th>Highest Temperature [°F] (HT)</th>
<th>Relative Humidity [%] (RH)</th>
<th>Maximum Wind Speed [m/s] (MWS)</th>
<th>Wind Gust [m/s] (WG)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

©2020 Mladen Kezunovic, All Rights Reserved
Prediction Results

Receiver Operating Characteristics Curve (ROC)

Event vs. Prediction of Failure

<table>
<thead>
<tr>
<th>Event</th>
<th>Failure (Y/N)</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>HT 86°F or below</td>
<td>Y=0</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>Y=1</td>
<td>47</td>
</tr>
<tr>
<td>HT 86°F - 89.6°F</td>
<td>Y=0</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>Y=1</td>
<td>47</td>
</tr>
<tr>
<td>HT 89.6°F or above</td>
<td>Y=0</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>Y=1</td>
<td>54</td>
</tr>
</tbody>
</table>
Conclusions

Big Data gives TRANSFORMATIONAL opportunities for risk assessment, management and mitigation

In the case of faults, we are able to PREDICT outages, which offers different opportunities for control, planning and protection mitigation actions

The outage PREDICTION can be at different spatiotemporal scales, minutes to hours, days and quarters, and as granular as a specific asset component

The time spent on DATA PREPARATION is the most consuming part of the overall development, and may take up to 80% of the development efforts

The TEST RESULTS are easy to confirm since the testing can be done using the data from the past when the outcomes of the events are already known
Thank you!

Questions?

Mladen Kezunovic
Tel: (979) 845-7509
E-mail: kezunov@ece.tamu.edu
Together - building a prosperous future

where energy is clean, abundant, reliable, safe, secure and affordable