Treatment Effect of Repairs to an Electrical Grid

Leveraging a Machine Learned Model of Structure Vulnerability

Rebecca Passonneau¹, Cynthia Rudin², Axinia Radeva¹, Ashish Tomar¹, Boyi Xie¹

¹. Center for Computational Learning Systems, Columbia University
². Sloan School of Management, MIT
Outline

• Introduction: low-voltage transmission (secondary) grid
• Secondary machine learning project
  • Data assembly and consolidation
  • Structure ranking (supervised machine learning)
  • Inspection program
• Methods: causal inference
• Results
• Conclusions
Consolidated Edison’s Secondary Electrical Grid

• A dense network of structures and cables provide power to NYC buildings

A manhole fire (rare) in the Village, April 2008 (Flickr)

Visualization of 2nd Ave & 83rd Street
• Manholes & service boxes (red is more vulnerable)
• Cables
Consolidated Edison’s Secondary Electrical Grid

High voltage transmission system

Secondary voltage electrical grid

Manhole

Service Box
Introduction

• Electrical structure inspection program
  • Consolidated Edison – utility company providing electrical power
  • Public Service Commission requires Con Edison to inspect all structures (e.g., manholes and service boxes)
  • Five-year cycle inspection program
Introduction

• Electrical structure inspection program
  • Consolidated Edison – utility company providing electrical power
  • Public Service Commission requires Con Edison to inspect all structures (e.g., manholes and service boxes)
  • Five-year cycle inspection program

• Program Assessment Goal
  • Identify which inspections have a measurable positive impact
  • Determine whether the structure vulnerability ranking facilitates the inspections analysis
Introduction

- Electrical structure inspection program
  - Consolidated Edison – utility company providing electrical power
  - Public Service Commission requires Con Edison to inspect all structures (e.g., manholes and service boxes)
  - Five-year cycle inspection program

- Program Assessment Goal
  - Identify which inspections have a measurable positive impact
  - Determine whether the structure vulnerability ranking facilitates the inspections analysis

- Method
  - Group structures based on vulnerability ranking
  - Causal analysis to evaluate inspection program outcome
Secondary Machine Learning Project

- **Raw data sources**
  - Structure information
    - Structure type (manhole, service box)
    - Unique identifier (asset id; or, 4-tuple of structure type, number, M&S plate)
    - Location (geo-coding)
    - Cover type (solid, gratings)
Secondary Machine Learning Project

- Raw data sources
  - Structure information
    - Structure type (manhole, service box)
    - Unique identifier (asset id; or, 4-tuple of structure type, number, M&S plate)
    - Location (geo-coding)
    - Cover type (solid, gratings)
  - Cable information
    - Number of main phase cables
    - Number of service phase cables
    - Connectivity: to/from structures
    - Age of cables
Secondary Machine Learning Project

• Raw data sources
  • Structure information
    • Structure type (manhole, service box)
    • Unique identifier (asset id; or, 4-tuple of structure type, number, M&S plate)
    • Location (geo-coding)
    • Cover type (solid, gratings)
  • Cable information
    • Number of main phase cables
    • Number of service phase cables
    • Connectivity: to/from structures
    • Age of cables
  • History of events
    • Stray voltage
    • “Burnouts” (low voltage; flickering lights; wire burnout; no lights area; …)
    • Smoking manhole
    • Fire, explosion
Free Text Data Source

- Emergency Control System (ECS) tickets
  - Noisy text; variable length; all data is unlabeled
  - Entries made over time by dispatch operators, onsite engineers

ECS (Emergency Control System) Trouble Tickets

1. 01/21/YR 18:45 FDNY-190 REPORTS A SMH STREET_1 & STREET_2
2. 01/21/YR 19:35 PERSON REPORTS THE TROUBLE HOLE IS SB-00001
3. 01/21/YR 19:45 FDNY-190 REPORTS A SMH STREET_1 & STREET_2
4. 01/21/YR 21:36 PERSON REPORTS IN SB-00001 BE FOUND 1 LEG
5. ON THE 5 WIRE NORTH BURNING IN THE STRUCTURE.....CUT/CLEAR
6. ED & RETIED SAME ..........................COMPLETE.............SS
7. ELIN REPT ADDED FOR INCIDENT:SMH 01/21/YR 22:02 BY PERSON.ID
8. REPORTED BY: FIRE DEPT
9. STRUC MSPLATE TYPE NUMBER COND COVTYPE? COVFOND DISTANCE
10. (1) MSPLATE ID SB 00001 W S Y 80
11. TYPE OF CURRENT: ALTERNATING CURRENT
12. VOLTAGE: 120/208V
13. APPROPRIATE SIZE: 500 MCM
14. CONDUCTOR (CODE: COPPER
15. POSSIBLE CAUSE OF THE INCIDENT: INSULATION BREAKDOWN
16. WEATHER CONDITIONS DURING THE INCIDENT: CLEAR
Free Text Data Source

- Emergency Control System (ECS) tickets
  - Noisy text; variable length; all data is unlabeled
  - Entries made over time by dispatch operators, onsite engineers

- Information Extraction (GATE – text engineering tool)
  - Filling the structured tables or templates with information extracted from textual documents
    - Structure type
    - Trouble hole (source)
    - Nature of problem
    - Other affected structures

---

ECS (Emergency Control System) Trouble Tickets

1. 01/21/YR 18:45 FDNY-190 REPORTS A SMH STREET_1 & STREET_2
2. 01/21/YR 19:35 PERSON REPORTS THE TROUBLE HOLE IS SB-00001
3. 01/21/YR 18:45 FDNY-190 REPORTS A SMH STREET_1 & STREET_2
4. 01/21/YR 19:35 PERSON REPORTS THE TROUBLE HOLE IS SB-00001
5. N/W/E STREET_1 & STREET_2...FOUND ON...SMOKING LIGHTLY
6. ON THE 5 WIRE NORTH BURNING IN THE STRUCTURE.......CUT/CLEAR
7. ED & RETIRED SAME ....................COMPLETE............SS
8. ELIN REPT ADDED FOR INCIDENT:SMH 01/21/YR 22:02 BY PERSON.ID
9. REPORTED BY: FIRE DEPT
10. STRUC MSPLATE TYPE NUMBER COND COVTYP COVFIND DISTANCE
11. (1) MSPLATE ID SB-00001 WA S Y 80
12. TYPE OF CURRENT: ALTERNATING CURRENT
13. VOLTAGE: 120/208V
14. APPROPRIATE SIZE: 500 MCM
15. CONDUCTOR COUPE: COPPER
16. POSSIBLE CAUSE OF THE INCIDENT: INSULATION BREAKDOWN
17. WEATHER CONDITIONS DURING THE INCIDENT: CLEAR
Free Text Data Source

- **Emergency Control System (ECS) tickets**
  - Noisy text; variable length; all data is unlabeled
  - Entries made over time by dispatch operators, onsite engineers

- **Information Extraction (GATE – text engineering tool)**
  - Filling the structured tables or templates with information extracted from textual documents
    - Structure type
    - Trouble hole (source)
    - Nature of problem
    - Other affected structures

### ECS (Emergency Control System) Trouble Tickets

<table>
<thead>
<tr>
<th>Ticket Labeling</th>
<th>Structure Labeling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ticket 1: 01/21/YR 18:45 FDNY-190 REPORTS A SMH STREET_1 &amp; STREET_2</td>
<td>Structure 1: STREET_1</td>
</tr>
<tr>
<td>Ticket 2: 01/21/YR 19:35 PERSON REPORTS THE TROUBLE HOLE IS SB-00001</td>
<td>Structure 2: STREET_2</td>
</tr>
<tr>
<td>Ticket 3: N/W/C STREET_1 &amp; STREET_2......FOUND ON ...SMOKING LIGHTLY</td>
<td>Structure 3: STREET_3</td>
</tr>
<tr>
<td>Ticket 4: 01/21/YR 21:36 PERSON REPORTS IN SB-00001 THE FOUND 1 LEG</td>
<td>Structure 4: STREET_4</td>
</tr>
<tr>
<td>Ticket 5: ON THE 5 WIRE NORTH BURNING IN THE STRUCTURE......CUT/CLEAR</td>
<td>Structure 5: STREET_5</td>
</tr>
<tr>
<td>Ticket 6: ED &amp; RETIED SAME ......................COMPLETE............SS</td>
<td>Structure 6: STREET_6</td>
</tr>
<tr>
<td>Ticket 7: LIN REPT ADDED FOR INCIDENT:SMH 01/21/YR 22:02 BY PERSON_ID</td>
<td>Structure 7: STREET_7</td>
</tr>
<tr>
<td>Ticket 8: REPORTED BY: FIRE DEPT</td>
<td>Structure 8: STREET_8</td>
</tr>
<tr>
<td>Ticket 9: STRUCTURE TYPE NUMBER COND COVTYP COVFIND DISTANCE</td>
<td>Structure 9: STREET_9</td>
</tr>
<tr>
<td>Ticket 10: (1) MSPLATE ID SB-00001 WA S Y 90</td>
<td>Structure 10: STREET_10</td>
</tr>
<tr>
<td>Ticket 11: TYPE OF CURRENT: ALTERNATING CURRENT</td>
<td>Structure 11: STREET_11</td>
</tr>
<tr>
<td>Ticket 12: VOLTAGE: 120/208V</td>
<td>Structure 12: STREET_12</td>
</tr>
<tr>
<td>Ticket 14: CONDUCTOR (CONE: COPPER</td>
<td>Structure 14: STREET_14</td>
</tr>
<tr>
<td>Ticket 15: POSSIBLE CAUSE OF THE INCIDENT: INSULATION BREAKDOWN</td>
<td>Structure 15: STREET_15</td>
</tr>
<tr>
<td>Ticket 16: WEATHER CONDITIONS DURING THE INCIDENT: CLEAR</td>
<td>Structure 16: STREET_16</td>
</tr>
</tbody>
</table>
Structure Ranking

• **Model training:** for 2009 prediction, train on labels from 2008, using features from 2007 and before

• **Label:** the structure was a source structure for a serious event during the prediction period.

• **Features**
  • Past events, cables, inspections
  • Selected using AUC values, information gain, backwards elimination

• **Ranking algorithm:** P-norm Push [Rudin2009] and related algorithms
  • Supervised bipartite ranking
  • Emphasizes the top of the ranked list
Structure Ranking

- Supervised bipartite ranking
Structure Ranking

The "P-Norm Push" Algorithm

\[
\min_{\lambda} F_p(\lambda) := \sum_{k=1}^{K} \left( \sum_{i=1}^{I} \exp(-f(x_i^+)+f(x_i^-)) \right)^p
\]

where \( f(x) = \sum_{j=1}^{n} \lambda_j h_j(x) \) and \( h_j : X \rightarrow [0,1], j = 1, \ldots, n \)

- Supervised bipartite ranking
Structure Ranking

\[ \min_{\lambda \in \mathbb{R}} F_p(\lambda) := \sum_{k=1}^{K} \left( \sum_{i=1}^{I} \exp(-f(\mathbf{x}_i^+ + f(\mathbf{x}_i^-)) \right)^p \]

where \( f(\mathbf{x}) = \sum_{j=1}^{n} \lambda_j h_j(\mathbf{x}) \) and \( h_j : X \rightarrow [0,1] \) \( j = 1, \ldots, n \)

**Supervised bipartite ranking**

**The objective is essentially a weighted version of AUC**
Structure Ranking

The "P-Norm Push" Algorithm

$$\min_{\lambda \in \mathbb{R}^n} F_p(\lambda) := \sum_{k=1}^{K} \left( \sum_{i=1}^{I} \exp(-f(x_i^+) + f(x_k^-)) \right)^p$$

where $f(x) = \sum_{j=1}^{n} \lambda_j h_j(x)$ and $h_j : \mathbb{R} \rightarrow [0,1]$  

- Supervised bipartite ranking
- The objective is essentially a weighted version of AUC

Scoring function and features:

$$\text{score}(\text{structure}) = \alpha_1 \times \text{Mentions} + \alpha_2 \times \text{RecentMentions} + \alpha_3 \times \text{TroubleHole} + \alpha_4 \times \text{RecentTrHole} + \alpha_5 \times \text{MainPhase} + \alpha_6 \times \text{ServPhase} + \alpha_7 \times \text{Serv_1960_1969}$$
Inspection Data

- Inspection reports
  - 126,478 reports in digital format from 2004 through 2009
  - There are 51,219 structures
  - A structure can have one or multiple inspections

<table>
<thead>
<tr>
<th>Number of inspections</th>
<th>Number of structures</th>
<th>Percentage of structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28,842</td>
<td>56.3111</td>
</tr>
<tr>
<td>2</td>
<td>10,226</td>
<td>19.9652</td>
</tr>
<tr>
<td>3</td>
<td>4,860</td>
<td>9.4887</td>
</tr>
<tr>
<td>4</td>
<td>2,661</td>
<td>5.1953</td>
</tr>
<tr>
<td>5</td>
<td>1,639</td>
<td>3.2000</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>30</td>
<td>8</td>
<td>0.0156</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>186</td>
<td>1</td>
<td>0.0020</td>
</tr>
</tbody>
</table>

- Inspection outcomes: clean; four categories (levels) of defect
**Inspection Data**

**Level 1**
- Cable contact
- Improper sealed ends
- Unsealed ducts

**Level 1 through 4**
Different priority

---

- cover
- duct
- service cables
- main cables
- crab
Inspection Data

Level 1
* Cable contact
* Improper sealed ends
* Unsealed ducts
...

Level 2
* Cover damage
* Damaged neutral wire
* Damaged main cables
...

Level 1 through 4
Different priority

high

low
Inspection Data

Level 1
* Cable contact
* Improper sealed ends
* Unsealed ducts
...

Level 1 through 4
Different priority

Level 2
* Cover damage
* Damaged neutral wire
* Damaged main cables
...

Level 3
* Structure damage ceiling floor
* Damaged secondary crabs
* Damaged secondary service cables
...

different priority
high
low

cover
duct
service cables
main cables
Inspection Data

Level 1
* Cable contact
* Improper sealed ends
* Unsealed ducts
...

Level 2
* Cover damage
* Damaged neutral wire
* Damaged main cables
...

Level 3
* Structure damage ceiling floor
* Damaged secondary crabs
* Damaged secondary service cables
...

Level 4 (not for repairs)
* Main cable replacement
* Service cable replacement
* Structure requires enlargement
...

Level 1 through 4
Different priority

Different priority:
- high
- low
Inspection Data

- **Level 1 through 4** (different priority of repair works)

- **Level distributions** (all possible outcomes)

<table>
<thead>
<tr>
<th>Result of inspections</th>
<th>Count (2008-2009)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean inspection</td>
<td>25,041</td>
<td>34.83</td>
</tr>
<tr>
<td>Level 1 only</td>
<td>17,928</td>
<td>24.94</td>
</tr>
<tr>
<td>Level 2 only</td>
<td>1,101</td>
<td>1.53</td>
</tr>
<tr>
<td>Level 3 only</td>
<td>18</td>
<td>0.03</td>
</tr>
<tr>
<td>Level 4 only</td>
<td>13,234</td>
<td>18.41</td>
</tr>
<tr>
<td>Level 1+2</td>
<td>1,417</td>
<td>1.97</td>
</tr>
<tr>
<td>Level 1+3</td>
<td>6</td>
<td>0.01</td>
</tr>
<tr>
<td>Level 1+4</td>
<td>9,127</td>
<td>12.70</td>
</tr>
<tr>
<td>Level 2+3</td>
<td>1</td>
<td>0.00</td>
</tr>
<tr>
<td>Level 2+4</td>
<td>1,652</td>
<td>2.30</td>
</tr>
<tr>
<td>Level 3+4</td>
<td>33</td>
<td>0.05</td>
</tr>
<tr>
<td>Level 1+2+3</td>
<td>2</td>
<td>0.00</td>
</tr>
<tr>
<td>Level 1+2+4</td>
<td>2,296</td>
<td>3.19</td>
</tr>
<tr>
<td>Level 1+3+4</td>
<td>17</td>
<td>0.02</td>
</tr>
<tr>
<td>Level 2+3+4</td>
<td>10</td>
<td>0.01</td>
</tr>
<tr>
<td>Level 1+2+3+4</td>
<td>7</td>
<td>0.01</td>
</tr>
<tr>
<td>Total</td>
<td>71,890</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Inspection Data

• Level 1 through 4 (different priority of repair works)

• Level distributions (all possible outcomes)

<table>
<thead>
<tr>
<th>Result of inspections</th>
<th>Count (2008-2009)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean inspection</td>
<td>25,041</td>
<td>34.83</td>
</tr>
<tr>
<td>Level 1 only</td>
<td>17,928</td>
<td>24.94</td>
</tr>
<tr>
<td>Level 2 only</td>
<td>1,101</td>
<td>1.53</td>
</tr>
<tr>
<td>Level 3 only</td>
<td>18</td>
<td>0.03</td>
</tr>
<tr>
<td>Level 4 only</td>
<td>13,234</td>
<td>18.41</td>
</tr>
<tr>
<td>Level 1+2</td>
<td>1,417</td>
<td>1.97</td>
</tr>
<tr>
<td>Level 1+3</td>
<td>6</td>
<td>0.01</td>
</tr>
<tr>
<td>Level 1+4</td>
<td>9,127</td>
<td>12.70</td>
</tr>
<tr>
<td>Level 2+3</td>
<td>1</td>
<td>0.00</td>
</tr>
<tr>
<td>Level 2+4</td>
<td>1,652</td>
<td>2.30</td>
</tr>
<tr>
<td>Level 3+4</td>
<td>33</td>
<td>0.05</td>
</tr>
<tr>
<td>Level 1+2+3</td>
<td>2</td>
<td>0.00</td>
</tr>
<tr>
<td>Level 1+2+4</td>
<td>2,296</td>
<td>3.19</td>
</tr>
<tr>
<td>Level 1+3+4</td>
<td>17</td>
<td>0.02</td>
</tr>
<tr>
<td>Level 2+3+4</td>
<td>10</td>
<td>0.01</td>
</tr>
<tr>
<td>Level 1+2+3+4</td>
<td>7</td>
<td>0.01</td>
</tr>
<tr>
<td>Total</td>
<td>71,890</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Most frequent
Inspection Data

• Level 1 through 4 (different priority of repair works)

• Level distributions (all possible outcomes)

<table>
<thead>
<tr>
<th>Result of inspections</th>
<th>Count (2008-2009)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean inspection</td>
<td>25,041</td>
<td>34.83</td>
</tr>
<tr>
<td>Level 1 only</td>
<td>17,928</td>
<td>24.94</td>
</tr>
<tr>
<td>Level 2 only</td>
<td>1,101</td>
<td>1.53</td>
</tr>
<tr>
<td>Level 3 only</td>
<td>18</td>
<td>0.03</td>
</tr>
<tr>
<td>Level 4 only</td>
<td>13,234</td>
<td>18.41</td>
</tr>
<tr>
<td>Level 1+2</td>
<td>1,417</td>
<td>1.97</td>
</tr>
<tr>
<td>Level 1+3</td>
<td>6</td>
<td>0.01</td>
</tr>
<tr>
<td>Level 1+4</td>
<td>9,127</td>
<td>12.70</td>
</tr>
<tr>
<td>Level 2+3</td>
<td>1</td>
<td>0.00</td>
</tr>
<tr>
<td>Level 2+4</td>
<td>1,652</td>
<td>2.30</td>
</tr>
<tr>
<td>Level 3+4</td>
<td>33</td>
<td>0.05</td>
</tr>
<tr>
<td>Level 1+2+3</td>
<td>2</td>
<td>0.00</td>
</tr>
<tr>
<td>Level 1+2+4</td>
<td>2,296</td>
<td>3.19</td>
</tr>
<tr>
<td>Level 1+3+4</td>
<td>17</td>
<td>0.02</td>
</tr>
<tr>
<td>Level 2+3+4</td>
<td>10</td>
<td>0.01</td>
</tr>
<tr>
<td>Level 1+2+3+4</td>
<td>7</td>
<td>0.01</td>
</tr>
<tr>
<td>Total</td>
<td>71,890</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Most frequent

The only guarantee repair works

Not repair works, but book-keeping
Causal Inference

• Question:
  • Do inspections result in a reduced incidence of events?
  • Level 1 inspection findings are addressed immediately
  • Level 2-3 inspection findings are deferred
  • Level 4 inspection findings are for internal bookkeeping (not for the PSC)
Causal Inference

• Question:
  • Do inspections result in a reduced incidence of events?
  • Level 1 inspection findings are addressed immediately
  • Level 2-3 inspection findings are deferred
  • Level 4 inspection findings are for internal bookkeeping (not for the PSC)

• Treatment to be investigated: Level 1 triggered repair
Causal Inference

• Question:
  • Do inspections result in a reduced incidence of events?
  • Level 1 inspection findings are addressed immediately
  • Level 2-3 inspection findings are deferred
  • Level 4 inspection findings are for internal bookkeeping (not for the PSC)

• Treatment to be investigated: Level 1 triggered repair

• To test for a treatment effect, find “twin” structures (matching)
  • Treated structure: $s_1$ received a Level 1 repair during time T
  • Control (untreated structure): $s_2$ (for all intents and purposes identical to $s_1$) did not receive any treatment (no inspections, clean or otherwise) during time T
  • Compare what happens to $s_1$ $s_2$ after time T
Relation of ranking to inspections

When do structures get inspected?

• **Ad hoc**: Whenever a crew enters a structure, e.g., in response to an event (e.g., no lights in area)

• **Targeted**: During last year or two of 5-year cycle, if no ad hoc inspections have occurred
Relation of ranking to inspections

When do structures get inspected?

• **Ad hoc:** Whenever a crew enters a structure, e.g., in response to an event (e.g., no lights in area)
• **Targeted:** During last year or two of 5-year cycle, if no ad hoc inspections have occurred

Structure ranking

• Reflects the likelihood that a structure will have an event the next year
• High ranked structures
• Are more likely to have ad hoc inspections
• Are more likely to have events
• Hypothesis: are more likely to have a measurable treatment effect, i.e., a reduced incidence of events
Define Treatment and Control Groups

- Solution: (post-)stratification
  - Divide data into distinct groups to increase homogeneity
  - Use rank to bin the groups
  - Check balance: do treatment/control have same range of values, e.g., rank
  - Check distribution: is there a similar change in number of treatment/control structures over the range of values?
Criteria

• Control
  • No inspections through 2008

• Treatment
  • No inspections before 2008, and in 2008, only structures with one Level 1 inspection

• The only difference: Level 1 repair in 2008
  • Thus, the control group is ensured to have no inspection triggered repairs until 2008
  • In 2008, the treatment group had Level 1 repairs only

• Eight strata (bins) defined by structure rank (based on feedback with Con Edison)
Results - Fisher's Exact Test to test for significance

<table>
<thead>
<tr>
<th>Cat</th>
<th>T Evt</th>
<th>C Evt</th>
<th>T Evt</th>
<th>C Evt</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>532</td>
<td>1152</td>
<td>6486</td>
<td>10981</td>
<td>3.055 × 10^-6</td>
</tr>
<tr>
<td>1</td>
<td>52</td>
<td>114</td>
<td>377</td>
<td>592</td>
<td>0.0370</td>
</tr>
<tr>
<td>2</td>
<td>88</td>
<td>152</td>
<td>690</td>
<td>1184</td>
<td>0.1067</td>
</tr>
<tr>
<td>3</td>
<td>105</td>
<td>191</td>
<td>841</td>
<td>1350</td>
<td>0.1830</td>
</tr>
<tr>
<td>4</td>
<td>92</td>
<td>210</td>
<td>969</td>
<td>1537</td>
<td>0.0030</td>
</tr>
<tr>
<td>5</td>
<td>72</td>
<td>172</td>
<td>1071</td>
<td>1752</td>
<td>0.0050</td>
</tr>
<tr>
<td>6</td>
<td>67</td>
<td>158</td>
<td>1161</td>
<td>1749</td>
<td>0.0036</td>
</tr>
<tr>
<td>7</td>
<td>53</td>
<td>114</td>
<td>1168</td>
<td>1847</td>
<td>0.0407</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>12</td>
<td>254</td>
<td>970</td>
<td>0.6203</td>
</tr>
</tbody>
</table>

- Treatment: + Events | - Events | Row Marginals
- Control: 114 | 592 | 706
- Column Marginals: 166 | 969 | 1135
- p=0.0370

Decreasing vulnerability
Conclusions

• Secondary electrical grid of Con Edison
• Trouble ticket processing (text engineering)
• Structure ranking (supervised bipartite ranking)
• Inspection program (quantifies the impact)
• Successfully applies data mining and machine learning methods to real world data
Thank you!
Some reference

• [http://www.osha.gov/SLTC/etools/electric_power/illustrated_glossary/substation_equipment/power_transformers.html](http://www.osha.gov/SLTC/etools/electric_power/illustrated_glossary/substation_equipment/power_transformers.html)

• Degradation as related to stray voltage and manhole events, final report.
