ADTRIBUTOR: REVENUE DEBUGGING IN ADVERTISING SYSTEMS

Ranjita Bhagwan, Rahul Kumar, Ram Ramjee, George Varghese
Surjyakanta Mohapatra, Hemanth Manoharan, Piyush Shah
MICROSOFT
ADVERTISING SYSTEMS ARE COMPLEX

Users

Publishers: bing.com, cnn.com, ...

Fraud Operators

Publisher Interfaces

Auctioning, Algorithms, ...

Advertiser Interfaces

Flowers Cars Insurance

Advertisers

Advertise Systems

Servers, Back-end storage (DB, ...)

Logging Component
REVENUE DEBUGGING IN ADVERTISING SYSTEMS

Why is Revenue/Revenue-per-search down anomalously?

- **Datacenter** in Dublin had latency issues that resulted in fewer ads being served.

- **Buckets 18, 23, and 24** were using a new algorithm for ad relevance that wasn’t working as expected.
  - Buckets: experimental trials with different algorithms

- The papal election was in progress, and users were searching for mainly non-monetizable *queries* such as “Pope”
CONTRIBUTIONS

1. Novel algorithm for root cause analysis in Ad Systems
   - Uses explanatory power, succinctness and surprise

2. Attribution for derived measures
   - E.g., attribute an element’s contribution to revenue-per-search (revenue/# searches)

3. Adtributor Tool
   - 95+% accuracy in identifying root causes in Ad Systems
   - Saves 1+ hour on average of manual troubleshooter time
OUTLINE

- Characteristics of Ad systems
- Root cause analysis
- Attribution for derived measures
- Adtributor Demo
- Evaluation
CHARACTERISTIC I: AGGREGATE ANALYSIS
CHARACTERISTIC II: FUNDAMENTAL AND DERIVED MEASURES

Searches

Ads per Search

Ads

Clicks per Ad

Clicks

Cost per Click

Revenue per Search

Revenue
CHARACTERISTIC III: MULTI DIMENSIONAL ANALYSIS
ROOT CAUSE ANALYSIS

- **Example**
  - Expected Revenue: $100, Actual Revenue: $80
  - Revenue down by 20% → anomaly!

- **Potential root causes**
  - One data center had $18 less revenue than forecasted
  - Three advertisers spent $20 less than forecasted
  - 10 buckets resulted in $20 less revenue than forecasted

- **Should we attribute root cause to dimension data center, advertiser or bucket? Which values?**
Explanatory power and succinctness

- Explanatory: root cause should explain most of change
- Succinctness: root cause likely to be few elements

- **DataCenter == X**
- **Advertiser == A1 OR Advertiser == A3 OR Advertiser == A4**

Pie charts show contribution to change by dimension-values.
• Root cause likely to **deviate most** from expectation
  • Relative entropy of actual vs expected probability (JS-divergence)

- **Advertiser == A1 OR Advertiser == A3 OR Advertiser == A4**
ALGORITHM

- Find the dimension and smallest set of values that maximally explain the anomalous change while also maximizing surprise

- Multi-objective optimization

- Greedy algorithm
  - Smallest set → each value contributes > 10% of change
  - Maximally explains → set should explain > 2/3 of change
  - Maximize surprise

```plaintext
Foreach m ∈ M // Compute surprise for all measures,
   Foreach i ∈ D // all dimensions,
      Foreach j ∈ E_i // all elements for a dimension
         p = V_{ij}^e(m)/V^e(m) // Equation 6
         q = V_{ij}^a(m)/V^a(m) // Equation 7
         S_{ij}(m) = D_{ij}(p, q) // Equation 10
      ExplanatorySet = {}
   Foreach i ∈ D
      SortedE = E_i.SortDescend(S_{ij}(m)) // Surprise
      Candidate = {}, Explains = 0, Surprise = 0
   Foreach j ∈ SortedE
      EP = (V_{ij}^a(m) - V_{ij}^e(m))/(V^a(m) - V^e(m))
      if (EP > T_{EEP}) // Occam’s razor
         Candidate.Add += E_{ij}
         Surprise += S_{ij}(m), Explains += EP
      if (Explains > T_{EEP}) // explanatory power
         Candidate.Surprise = Surprise,
         ExplanatorySet += Candidate, break
   //Sort Explanatoryset by Candidate.Surprise
   Final = ExplanatorySet.SortDescend(Surprise)
Return Final.Take(3) // Top 3 most surprising
```
### Why derived measures?

**Intuition:** use expected value for all other elements and actual values for only this element

- Captured by Partial Derivatives in Finite Difference Calculus

\[
\frac{F(.)}{G(.)} = \frac{(\Delta F \cdot G - \Delta G \cdot F)}{(G \cdot (G + \Delta G))}
\]

### How do we attribute for derived measures?

**Intuition:** use expected value for all other elements and actual values for only this element

### Above 20% threshold

<table>
<thead>
<tr>
<th>Advertiser</th>
<th>Estimated Revenue</th>
<th>Actual Revenue</th>
<th>% change</th>
<th>Advertiser</th>
<th>Estimated Clicks</th>
<th>Actual Clicks</th>
<th>% change</th>
<th>Advertiser</th>
<th>Estimated CPC</th>
<th>Actual CPC</th>
<th>% change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>100</td>
<td>90</td>
<td>-10</td>
<td>Overall</td>
<td>500</td>
<td>580</td>
<td>16</td>
<td>Overall</td>
<td>0.2</td>
<td>0.155</td>
<td>-22.5</td>
</tr>
<tr>
<td>A1</td>
<td>50</td>
<td>10</td>
<td>400</td>
<td>A1</td>
<td>100</td>
<td>20</td>
<td>-100</td>
<td>A1</td>
<td>0.5</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>A2</td>
<td>200</td>
<td>360</td>
<td>200</td>
<td>A2</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>40</td>
<td>70</td>
<td>-300</td>
<td>A3</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>A3</td>
<td>0.4</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>A4</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>A4</td>
<td>0.1</td>
<td>0.1</td>
<td></td>
</tr>
</tbody>
</table>
DEMO

Adtributor
**EVALUATION**

- Evaluated 128 alerts generated over a 2 week period over 8 markets (US, UK, DE, FR: PC, Mobile for each)
- Compared Adtributor output with manual root-causing
- Time saved: 1+ hour on average per alert

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anomalies</td>
<td>128</td>
</tr>
<tr>
<td>No. of matches</td>
<td>118</td>
</tr>
<tr>
<td>Manual errors</td>
<td>4</td>
</tr>
<tr>
<td>Adtributor’s errors</td>
<td>5</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>1</td>
</tr>
<tr>
<td>Accuracy</td>
<td>95.3%</td>
</tr>
</tbody>
</table>
# RELATED WORK

<table>
<thead>
<tr>
<th></th>
<th>Root causing</th>
<th>Multiple Dimensions</th>
<th>Derived Measures</th>
</tr>
</thead>
</table>
| **Network Component**  
*Failure Isolation (e.g., SCORE, Sherlock, etc.)* | Explanatory Power, Succinctness | Does not handle                                           | Does not handle |
| **Network Traffic**  
*Pattern Finding (Autofocus, HHH)* | Explanatory Power, Succinctness | Explores all combinations of dimensions *dynamically*, Heuristic: unexpectedness | Does not handle |
| **Data mining**  
* (Summarization, Surprising Patterns) | Explanatory Power, Succinctness | Many techniques (e.g., Minimize description length) | Does not handle |
| **Revenue Debugging** | Explanatory Power, Succinctness | Explores single dimensions Pre-declared *statically*  
Surprise: JS divergence | Partial derivative, Finite differences |
Algorithm for Root Cause Analysis in Advertising Systems
- Uses explanatory power, succinctness, and surprise

Attribution for derived measures
- Finite difference, partial derivative-based approach

Adtributor tool
- 95+% accuracy, saves 1+ hour of manual troubleshooting time
APPLYING OUR APPROACH MORE GENERALLY

- This problem/solution is not specific to advertising
- Datacenter Diagnostics problem (Bodik et al., Eurosys 2010)
  - Problem: When there is a slowdown in the datacenter, where is the slowdown? Is it CPU, Memory or Disk that is the bottleneck?
- Derived metric attribution
  - MoS score attribution in VOIP networks: which link is responsible for drop in the Mean Opinion Score (MoS) for a given VOIP call?
CASE STUDY: ANOMALOUS REVENUE DROP

Dimension: Browser
- Expected
  - B1: 12%
  - B2: 12%
  - B3: 16%
  - B4: 60%
- Actual
  - B1: 8%
  - B2: 18%
  - B3: 74%

Dimension: Bucket
- Expected
  - B1
  - B2
- Actual
  - B1
  - B2

Dimension: Data Center
- Expected
  - DC 1
  - DC 2
  - DC 3
  - DC 4
- Actual
  - DC 1
  - DC 2
  - DC 3

➤ Maximum surprise (deviation from expected value) seen for the browser dimension
- Configuration error caused no ads to be shown on B3 for that time