# ADTRIBUTOR: REVENUE DEBUGGING IN ADVERTISING SYSTEMS
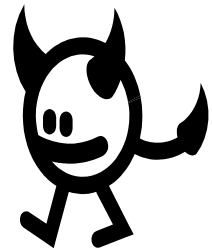
Ranjita Bhagwan, Rahul Kumar, Ram Ramjee, George Varghese

Surjyakanta Mohapatra, Hemanth Manoharan, Piyush Shah
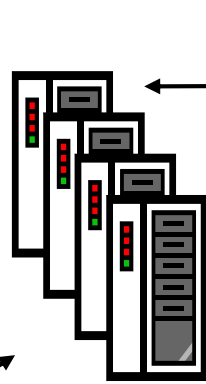
MICROSOFT

# ADVERTISING SYSTEMS ARE COMPLEX



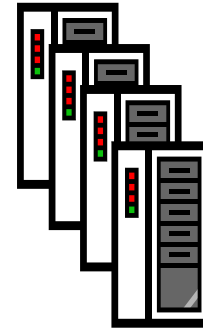Users

Advertising System

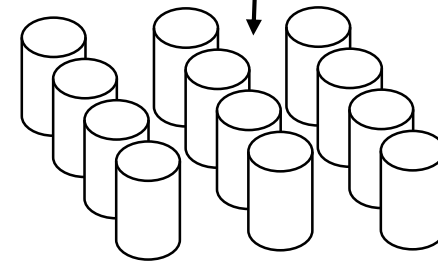Publisher Interfaces

Auctioning, Algorithms, …

Advertiser Interfaces

Servers, Back-end storage (DB, …)

Logging Component

Publishers: bing.com, cnn.com, …

| Flowers | Cars | Insurance |

Fraud Operators

Advertisers

# REVENUE DEBUGGING IN ADVERTISING SYSTEMS

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

*Measures*

*Derived Measures*

| Anomaly Detection Problem | → ← | Root Cause Analysis |

*Dimensions*

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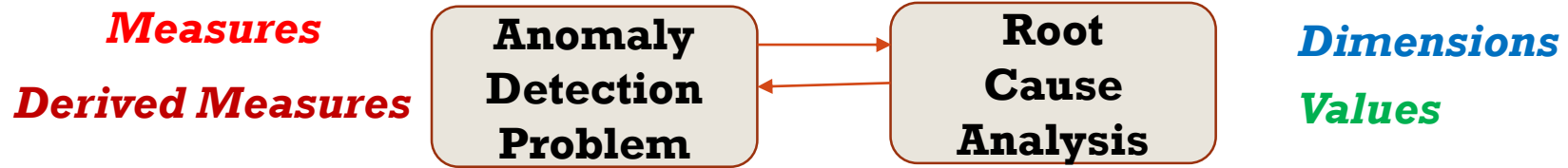
- **Datacenter** in **Dublin** had latency issues that resulted in fewer ads being served

- **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"**

3

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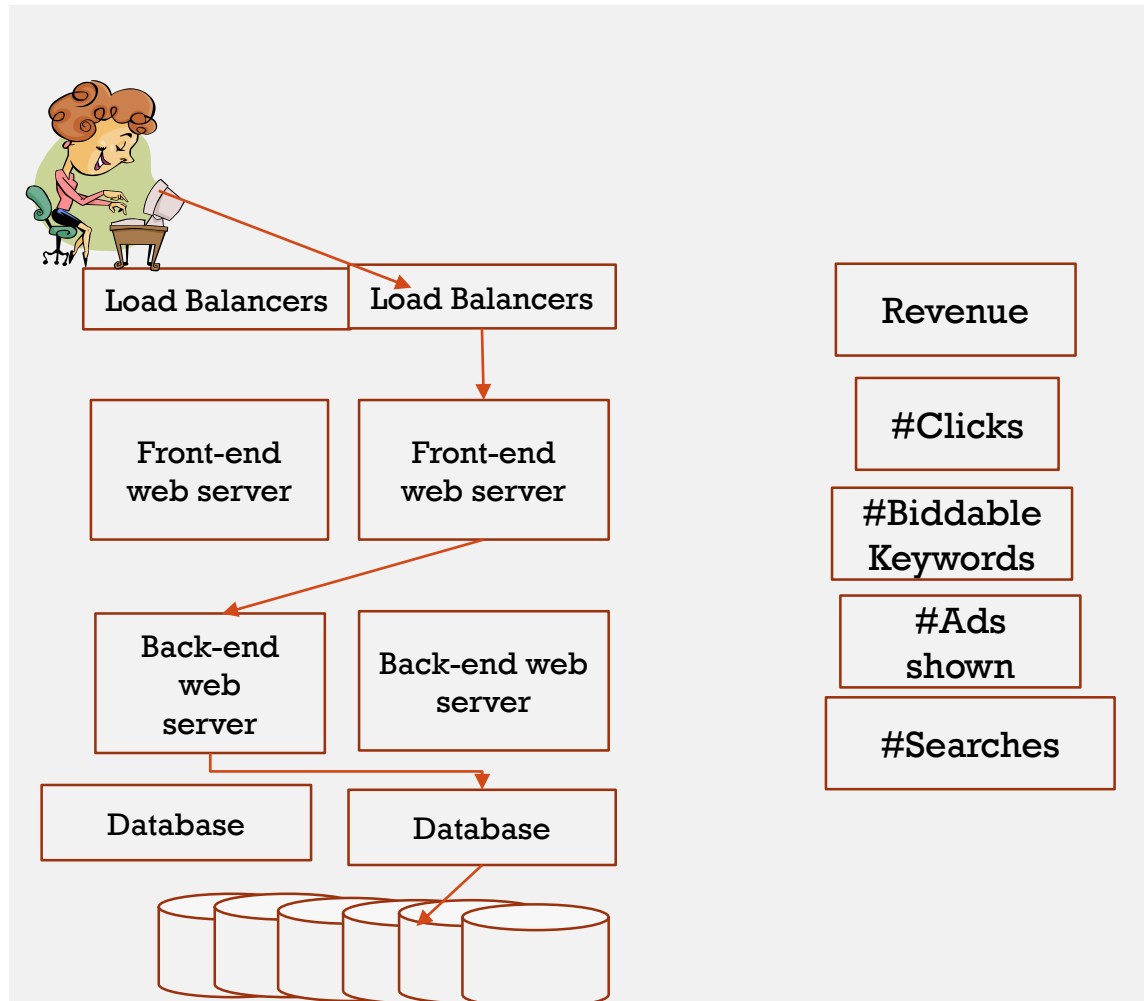
3. **Adtributor Tool**
   - 95+% accuracy in identifying root causes in Ad Systems
   - Saves 1+ hour on average of manual troubleshooter time
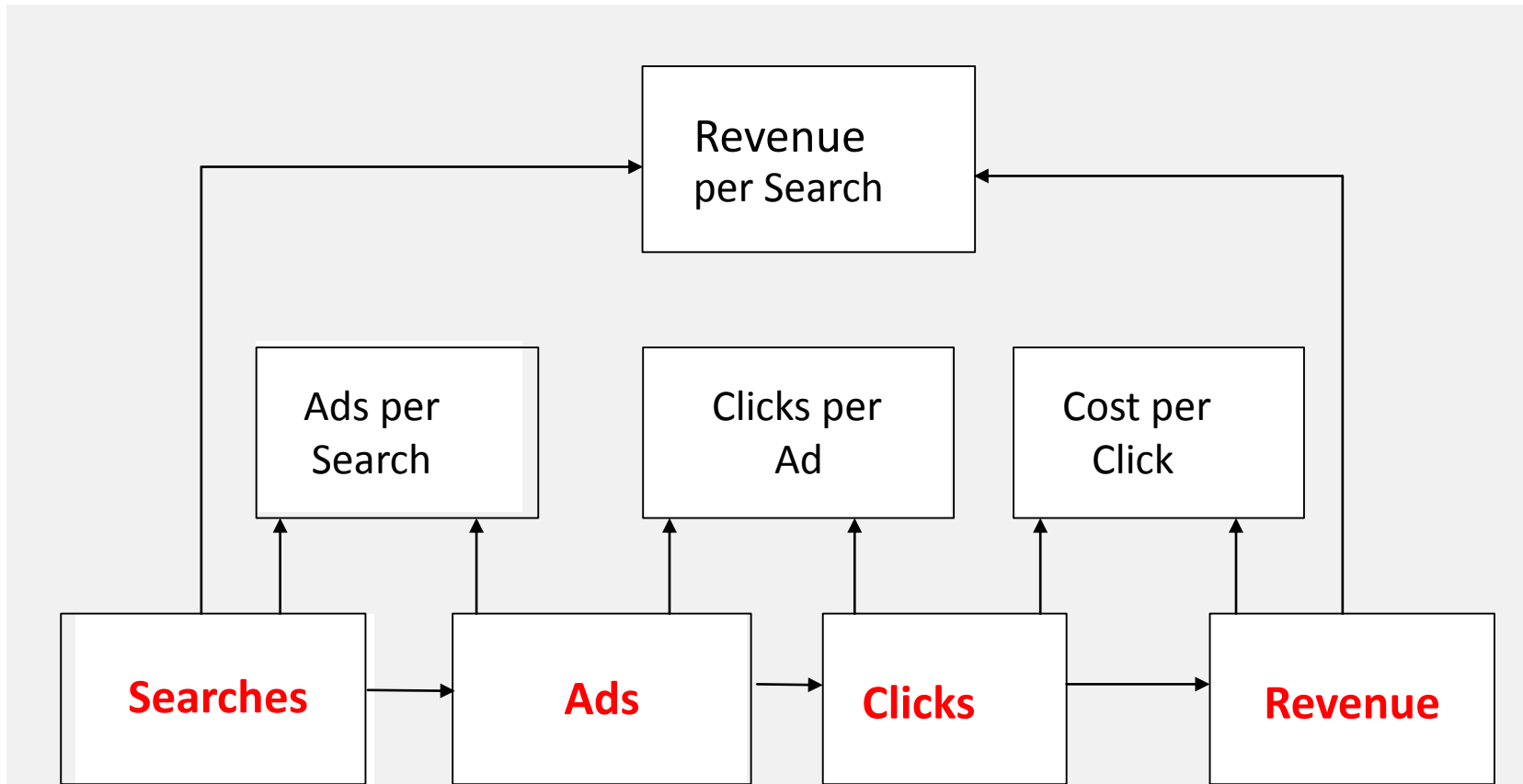
# OUTLINE

- Characteristics of Ad systems

- Root cause analysis

- Attribution for derived measures

- Adtributor Demo

- Evaluation

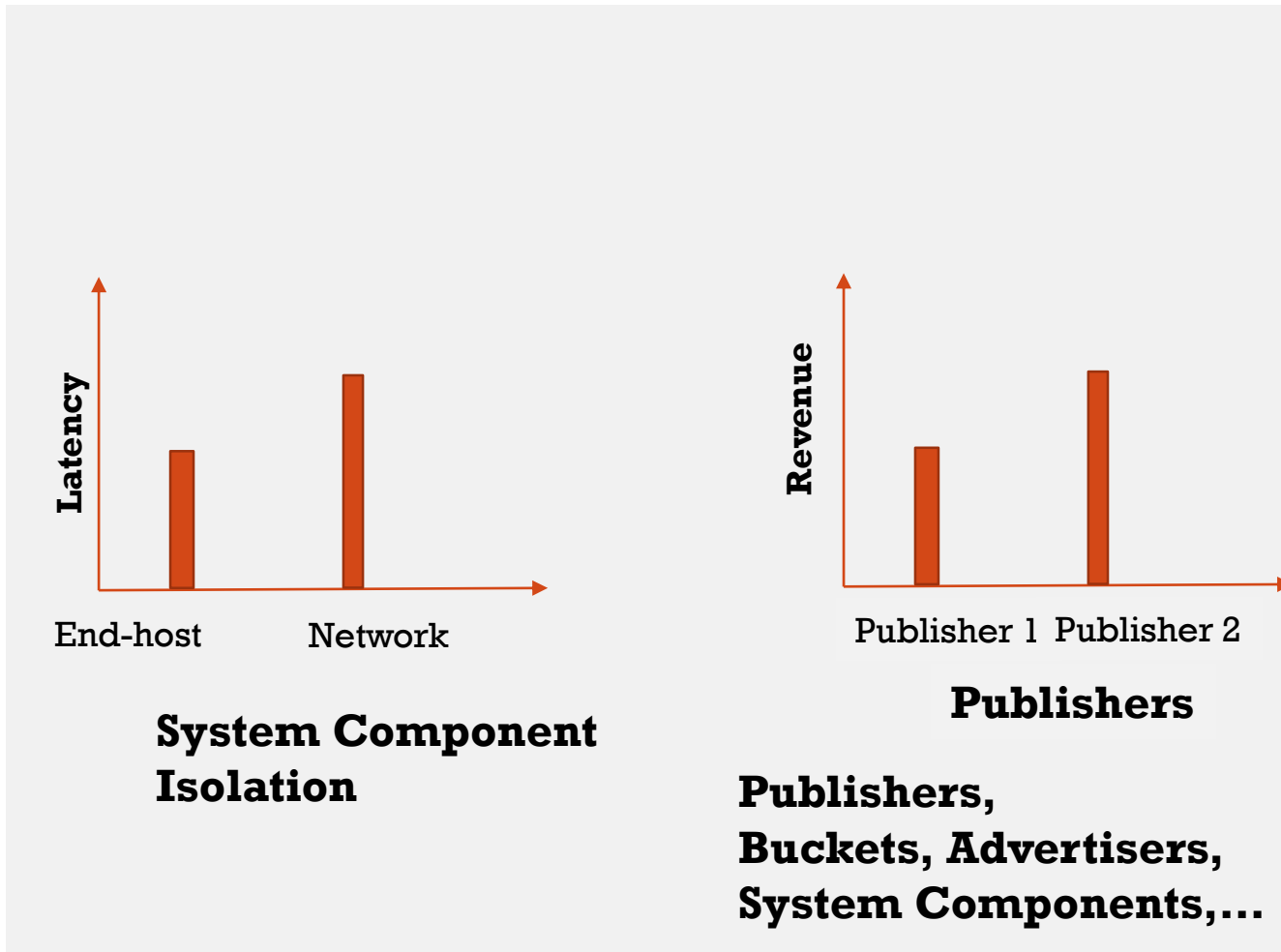# CHARACTERISTIC I: AGGREGATE ANALYSIS

# CHARACTERISTIC II: FUNDAMENTAL AND DERIVED MEASURES

# CHARACTERISTIC III: MULTI DIMENSIONAL ANALYSIS



**System Component Isolation**

**Publishers**

**Publishers, Buckets, Advertisers, System Components,...**

# ROOT CAUSE ANALYSIS

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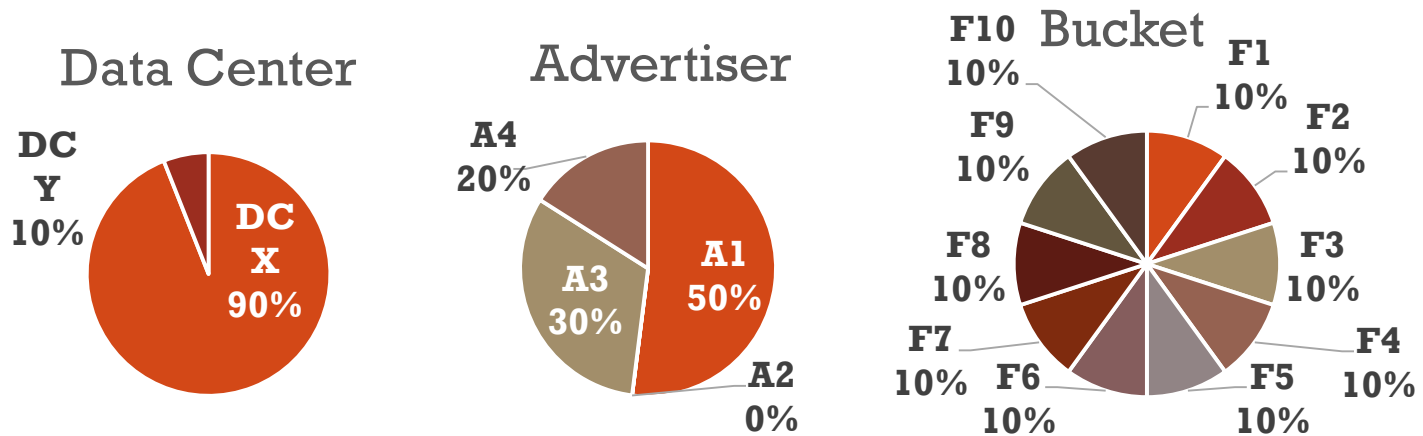  - Expected Revenue: $100, Actual Revenue: $80
  - Revenue down by 20% → anomaly!

- 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

**Data Center**

- DC Y 10%
- DC X 90%

**Advertiser**

- A4 20%
- A3 30%
- A1 50%
- A2 0%

**Bucket**

- F10 10%
- F1 10%
- F2 10%
- F3 10%
- F4 10%
- F5 10%
- F6 10%
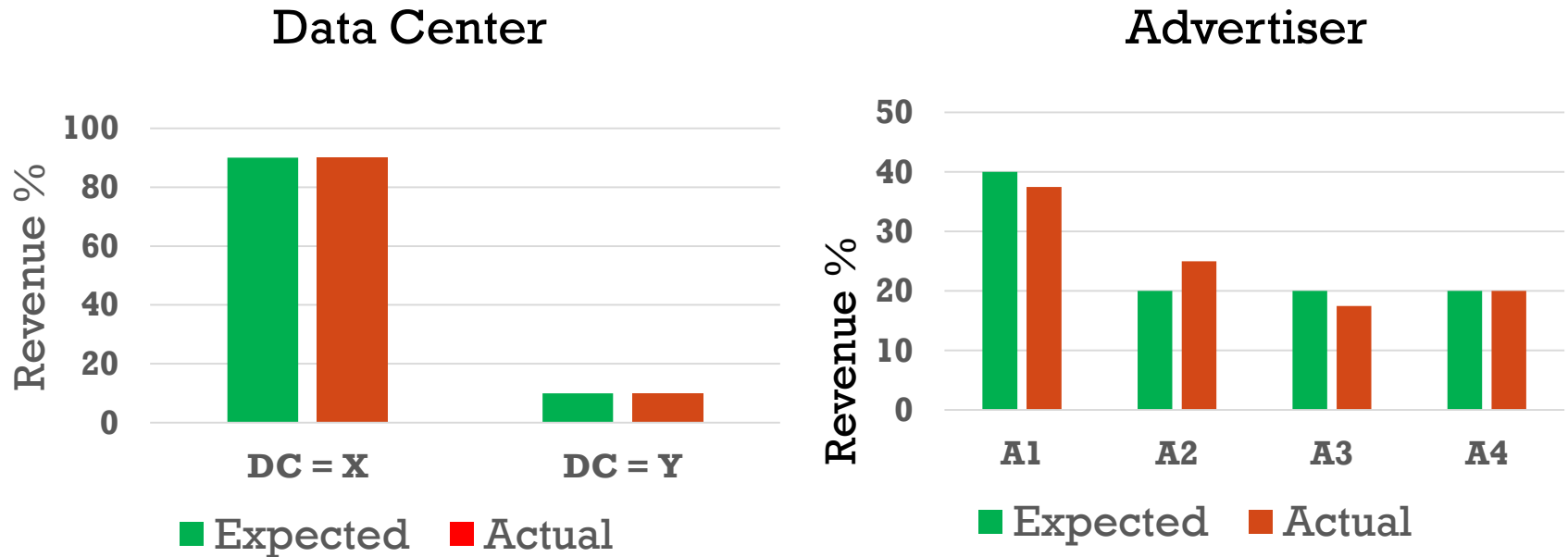- F7 10%
- F8 10%
- F9 10%

Pie charts show contribution to change by dimension-values.

- Explanatory: root cause should <span style="color:red">explain most</span> of change
- Succinctness: root cause likely to be <span style="color:red">few elements</span>

➢ **DataCenter == X**

➢ **Advertiser == A1 OR Advertiser == A3 OR Advertiser == A4**

# SURPRISE

## Data Center



## Advertiser



- 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 <span style="color:red">dimension and smallest set of values</span> that <span style="color:red">maximally explain the anomalous change</span> while also <span style="color:red">maximizing surprise</span>

- Multi-objective optimization

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

```
1  Foreach m ∈ M // Compute surprise for all measures,
2      Foreach i ∈ D // all dimensions,
3          Foreach j ∈ E_i // all elements for a dimension
4              p = V_{ij}^e(m)/V^e(m) // Equation 6
5              q = V_{ij}^a(m)/V^a(m) // Equation 7
6              S_{ij}(m) = D_{JS}(p,q) // Equation 10
7  ExplanatorySet = {}
8  Foreach i ∈ D
9      SortedE = E_i.SortDescend(S_{ij}(m)) //Surprise
10     Candidate = {}, Explains = 0, Surprise = 0
11     Foreach j ∈ SortedE
12         EP = (V_{ij}^a(m) − V_{ij}^e(m))/(V^a(m) − V^e(m))
13         if (EP > T_{EEP}) // Occam's razor
14             Candidate.Add += E_{ij}
15             Surprise += S_{ij}(m), Explains += EP
16         if (Explains > T_{EP}) // explanatory power
17             Candidate.Surprise = Surprise,
18             ExplanatorySet += Candidate, break
19 //Sort Explanatoryset by Candidate.Surprise
20 Final = ExplanatorySet.SortDescend(Surprise)
21 Return Final.Take(3) // Top 3 most surprising
```

# ATTRIBUTION FOR DERIVED MEASURES

Explanatory Power of element j in dimension i for measure m is simply=(A_ij(m) - F_ij(m)))/(A(m) - F(m)). e.g. A1: (10-50)/(90-100)=400%

- ## Why derived measures?

**Below 20% threshold**   **Below 20% threshold**   **Above 20% threshold**

| Adver-tisr | Estimated Revenue | Actual Revenue | % change | Adver-tisr | Estimated Clicks | Actual Clicks | % change | Adver-tisr | Estimated CPC | Actual CPC | % change |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Overall | 100 | 90 | -10 | Overall | 500 | 580 | 16 | Overall | 0.2 | 0.155 | -22.5 |
| A1 | 50 | 10 | 400 | A1 | 100 | 20 | -100 | A1 | 0.5 | 0.5 | |
| A2 | 0 | 0 | 0 | A2 | 200 | 360 | 200 | A2 | 0 | 0 | |
| A3 | 40 | 70 | -300 | A3 | 100 | 100 | 0 | A3 | 0.4 | 0.7 | |
| A4 | 10 | 10 | 0 | A4 | 100 | 100 | 0 | A4 | 0.1 | 0.1 | |

- ## How do we attribute for derived measures?

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

  e.g. A1: (10+0+40+10)/(20+200+100+100) = 0.143  , (0.2-0.143)/0.2=-28.5%
  A2: (0+50+40+10)/(360+100+100+100)=0.152,  (0.2-0.152)/0.2=-24%

- ## Captured by Partial Derivatives in Finite Difference Calculus

  F(.)/G(.) = (Delta_F*G – Delta_G*F)/(G * (G + Delta_G))

13

# DEMO

[Adtributor](Adtributor)

14

# EVALUATION

| Parameter | Value |
|---|---|
| Anomalies | 128 |
| No. of matches | 118 |
| Manual errors | 4 |
| Adtributor's errors | 5 |
| Ambiguous | 1 |
| Accuracy | 95.3% |

- 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

# RELATED WORK

| | Root causing | Multiple Dimensions | Derived Measures |
|---|---|---|---|
| **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 |

# SUMMARY
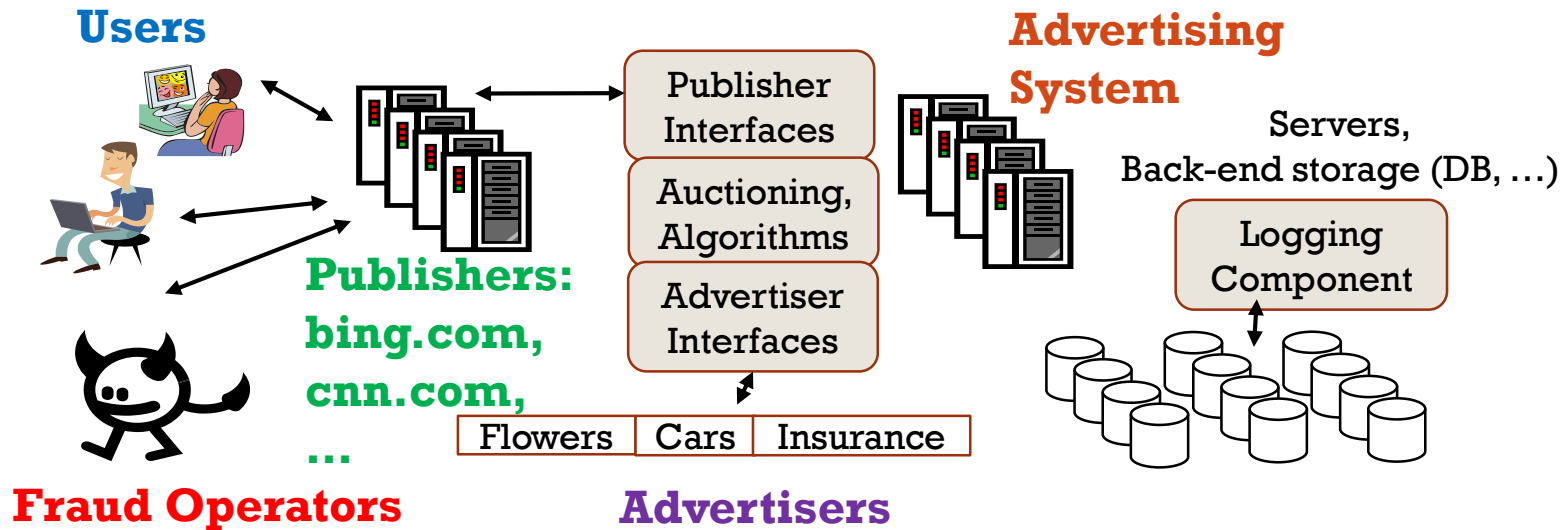


- ➢ Algorithm for Root Cause Analysis in Advertising Systems
  - • Uses explanatory power, succinctness, and surprise

- ➢ 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

Actual



- B1
- B2
- B3
- B4

- B1
- B2
- B3
- B4

**Dimension: Bucket**

Expected

Actual



**Dimension: Data Center**

Expected

Actual



- DC 1
- DC 2
- DC 3
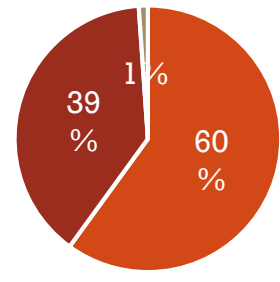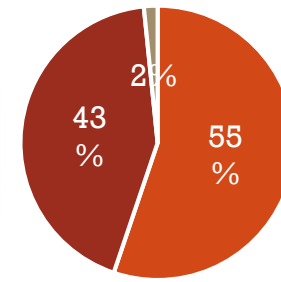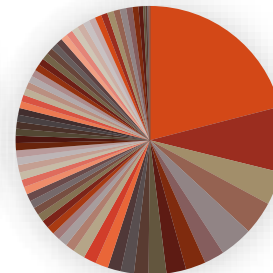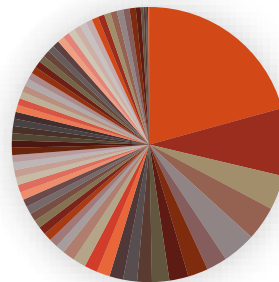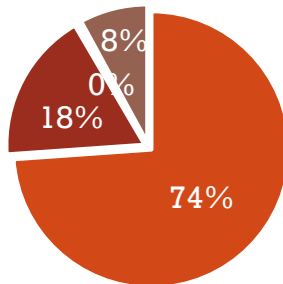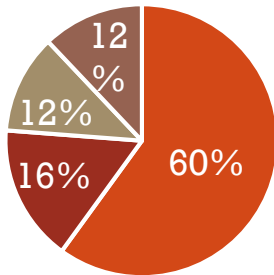
- DC 1
- DC 2
- DC 3

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

19