Digital Attribution

A practical approach to enhanced digital media performance
Agenda

• Overview: What you first need to know about the data
  – Opportunity and challenges of digital measurement
  – Digital media ecosystem
  – Digital media data primer

• How to Measure: Testing vs. Modeling
  – Testing Applications
  – Attribution Modeling in Action
Effective Digital Measurement

- There is **no perfect formula** for digital marketing programs

- In spite of challenges there is huge room for optimization beyond what is being done today

- General principles to effective digital measurement
  1. Understand the data and its limitations
  2. Connect the data
  3. Use a combination of testing and modeling

"Now that desk looks better. Everything's squared away, yessir, squaaaaaaaared away."

© 2012 Merkle Inc. All Rights Reserved | Confidential
Wouldn’t it be Nice if the Data Looked Like This?

BACK IN 2003, WE HAD SOMETHING CALLED BOREDOM.

Saw TV ads on Jan 1, 2 and 4
Received DM on Jan 10
Saw online Display ad on Jan 9, 10, and 12
Talked with friend at concert about product on Jan 12
Watched add with YouTube video on Jan 16
Paid search on Jan 16
Purchased product on Jan 16
Reality: Fragmented Media
Landscape: Fragmented Data
This Doesn’t Help Either

Traditional Funnel

Reality
Digital Data 101

Between your ad server and enterprise web analytics platform (sorry Google analytics) you have most of the data you need.

<table>
<thead>
<tr>
<th>Media</th>
<th>Primary Source</th>
<th>Primary Keys</th>
<th>What you can get</th>
</tr>
</thead>
<tbody>
<tr>
<td>Website (first party)</td>
<td>Web Analytics Platform (Ad server-secondary)</td>
<td>Cookie/ IP</td>
<td>• Detailed site engagement (view/click/searches/downloads etc.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Referral (How you got to site)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Campaign data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Form data</td>
</tr>
<tr>
<td>Search</td>
<td>Web Analytics Platform/Ad Server</td>
<td>Cookie/ Keyword</td>
<td>• Search clicks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Keyword/campaign data</td>
</tr>
<tr>
<td>Display, video, mobile display</td>
<td>Ad Server/ Demand side platform (DSP)</td>
<td>Cookie/ IP</td>
<td>• Ad impressions</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Ad clicks</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Ad context (site, url)</td>
</tr>
<tr>
<td>Mobile SMS text</td>
<td>Mobile platform</td>
<td>Mobile Phone #/ email</td>
<td>• SMS send</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• SMS click</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Ad meta data</td>
</tr>
<tr>
<td>Social</td>
<td>Facebook Open Graph/ Twitter/ etc</td>
<td>Social handle, email</td>
<td>• Posts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Likes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Friends data</td>
</tr>
</tbody>
</table>
Data Needs to be Linked to Unlock Value

Anonymous site visitors
- Requests information
- Fills out application
- Opens account
- Logs into account

Organic Search
- Web Site

Paid Search
- Landing Page

Social (FB, Youtube)
- Landing Page

Online Display/Video
- Landing Page

Email/SMS
- Landing Page

DM
- Landing Page

Identified visitors
- John Smith
  - Sent DM piece on 11/18
  - Saw display ad on 12/2
  - Searched for ‘Auto loan’ on 12/2 and clicked on paid search ad
  - Initiated loan online on 12/4
  - Completed loan at brand on 12/15

Customer Database
How Long do Cookies Really Persist?

Cookie deletion is a challenge. It needs to be factored into measurement methodologies

- 2011 Comscore study sited that 33% of users delete first party cookies in a given month. Nearly 43% of users delete third party cookies in a month.

- Cookie deletion plus the use of multiple devices and browsers means that audience reports are often inflated by 2.7X
What About IP Addresses?

Forget about neatly merging and reporting all of your offline and online media interactions at the Zip Code level

- It is harder to find published data IP uniqueness. Most of what I have learned has come from confidential communications with IP data providers and demand side platform vendors (DSPs) as well as our own analysis.

- About 85-90% of US IP addresses can be accurately tracked back to a DMA

- About 60% of devices with an IP address can be traced back to a known SCF and about 45% to the zip level

- Less than 25% of IPs can be reliably tracked back to a residence over at least one month’s time
There are a lot of data challenges when working with digital media but we still have options.
How to Measure

I didn't have any accurate numbers so I just made up this one.

Studies have shown that accurate numbers aren't any more useful than the ones you make up.

How many studies showed that?

Eighty-seven.
Better Digital Measurement Starts with Moving Beyond the Status Quo

Look Familiar?

Prior Media Interactions

Last Click

Closer to Reality

View Thru 2 5%
View Thru 1 15%
2nd Click 20%
Most Recent Click 60%

But, we’re still missing a few things

- Data makes it look like we have more unique visitors than we actually do (cookie deletion/multiple devices)
- Have not accounted for mass media and offline interactions
Measurement Methodologies

“The Truth” (Nirvana)

Last Click Approach (Poor)

Direct Rules Based (Better)

Incremental Approach (Best!)

Traditional Attribution Methodology Flaws
- Over-attribute to bottom of funnel media
- Ignore brand & offline media impact
- Weighting should not be arbitrary

We can do better with Testing & Modeling

Offline
- Direct Mail Sent
- Call Center

Digital
- Paid Search Click
- Website Visit
- Social visit

Estimated exposure

<table>
<thead>
<tr>
<th>Media Type</th>
<th>Offline</th>
<th>Direct</th>
<th>TV View</th>
<th>Newspaper View</th>
<th>Display View</th>
<th>Paid Search Click</th>
<th>Website Visit</th>
<th>Social visit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call Center</td>
<td>M</td>
<td>T</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct Mail</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Media Type</th>
<th>Estimated exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>TV View</td>
<td></td>
</tr>
<tr>
<td>Display View</td>
<td></td>
</tr>
<tr>
<td>Paid Search Click</td>
<td></td>
</tr>
<tr>
<td>Website Visit</td>
<td></td>
</tr>
<tr>
<td>Social visit</td>
<td></td>
</tr>
</tbody>
</table>
Two Primary Methods: Controlled testing vs. Attribution modeling

<table>
<thead>
<tr>
<th>Requirements</th>
<th>Testing</th>
<th>Modeling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lots of media</td>
<td>Weak</td>
<td>Strong</td>
</tr>
<tr>
<td>Lots of interactions</td>
<td>Weak</td>
<td>Strong</td>
</tr>
<tr>
<td>Indirect Media</td>
<td>Weak</td>
<td>Strong</td>
</tr>
<tr>
<td>Inability to impact media plan for testing</td>
<td>Weak</td>
<td>Strong</td>
</tr>
<tr>
<td>Low spend variability</td>
<td>Strong</td>
<td>Weak</td>
</tr>
<tr>
<td>Direct Media (ex. DM)</td>
<td>Strong</td>
<td>Weak</td>
</tr>
<tr>
<td>Small spend media</td>
<td>Strong</td>
<td>Weak</td>
</tr>
<tr>
<td>High Precision</td>
<td>Strong</td>
<td>Weak</td>
</tr>
<tr>
<td>Limited data management infrastructure</td>
<td>Strong</td>
<td>Weak</td>
</tr>
<tr>
<td>Digital Media Type</td>
<td>Most Accurate Measurement Level</td>
<td>Holdout Based Testing (Y/N)</td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Email</td>
<td>Customer</td>
<td>YES</td>
</tr>
<tr>
<td>Mobile SMS</td>
<td>Customer</td>
<td>YES</td>
</tr>
<tr>
<td>Website</td>
<td>Cookie</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Display</td>
<td>Cookie</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paid Search</td>
<td>Cookie</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affiliate/Lead Gen</td>
<td>Cookie</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobile App</td>
<td>App</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Media</td>
<td>Source</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example 1: Display Incremental

Goal: Estimate incremental impact of display campaigns

Situation:
• Difficult to justify display spend on click thru rates
• Don’t trust that view-thru conversions are incremental

Solution:
• Placebo testing
• Control get’s public service ads only
• Test get’s normal display creative

Outcome:
• Conservative estimate for display impact
• Most of the impact within 24 hours

Control: PSA

Test: Campaign

- Banner add click thru rates are only .09% (Industry wide)!
- View thru revenue 3x click thru revenue
Example 2: Paid vs. Organic Search

**Goal:** Estimate Incremental impact of paid search terms

**Situation:**
- High % of spend on branded paid search
- How much is really incremental?

**Solution**
- Branded keyword blackout test
- Google trend data used to estimate organic orders w/o paid search
- Cannibalized orders = Est. Organic orders – actual organic orders

**Outcome**
- Estimate 40% cannibalization between Paid and organic search
- Adjust keyword values accordingly in targeting

Shut off SEM

Compare to post period

Model established while paid search off

Est. organic orders w/o paid search
Example 3: DM + Display vs. DM alone

**Goal:** Estimate Incremental performance for each digital campaign

**Situation:**
- Financial Services Company
- Large DM program
- Wanted to see if Display could lift DM

**Solution:**
- Controlled test
- Offline to online direct match
- Targeted display and DM

**Outcome:**
- Able to measure display lift on DM program
- 17% lift on DM with display

1. **Pull Prospect Lists**
   - Database
   - DM Only
   - DM + Display
   - Display Only

2. **Target in media**
   - RTB Inventory
   - Display Responsive Audiences
   - DM Responsive Audiences

3. **Analyze Results**
   - 95% of US Population
   - Auction Marketplace
How to Measure

Attribution Modeling
Attrition Modeling Example

- **Challenge:** Inaccurate digital performance metrics make it difficult to optimize digital spend
  - KPI assumptions are very sensitive to attribution assumptions: Cost per inquiry estimates vary by more than 30% based on calculation assumptions
  - Out of the box estimation techniques are inadequate
- **Solution:**
  - Automated, model-driven solution that produces more accurate KPI estimates for digital from a media level down to the granular placement.
Solution: Bottom up and Top Down Modeling

**Top-down media mix model**
*(Traditional media mix model: DMA by week level, 12+ months of data)*

Cost per inquiry by tactic

- **National media (TV & radio)**: $140
- **Local media (TV & radio)**: $200
- **Direct mail**: $180
- **Digital**: $83

**Bottom-up digital attribution model**
*(New media mix model: zip by day level, can be built with only 1 month of data, 70+ programs estimated)*

Cost per inquiry by tactic

<table>
<thead>
<tr>
<th>Tactic</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display</td>
<td>$60</td>
</tr>
<tr>
<td>Search</td>
<td>$91</td>
</tr>
<tr>
<td>Search 1</td>
<td>$115</td>
</tr>
<tr>
<td>Search 2</td>
<td>$87</td>
</tr>
<tr>
<td>Search 3</td>
<td>$39</td>
</tr>
<tr>
<td>Search 1 - LG 1</td>
<td>$153</td>
</tr>
<tr>
<td>Search 2 - LG 2</td>
<td>$53</td>
</tr>
<tr>
<td>Search 3 - LG 3</td>
<td>$39</td>
</tr>
<tr>
<td>LG 1</td>
<td>$190</td>
</tr>
<tr>
<td>LG 2</td>
<td>$163</td>
</tr>
<tr>
<td>LG 3</td>
<td>$87</td>
</tr>
<tr>
<td>LG 4</td>
<td>$74</td>
</tr>
<tr>
<td>LG 5</td>
<td>$32</td>
</tr>
<tr>
<td>LG 6</td>
<td>$29</td>
</tr>
<tr>
<td>LG 7</td>
<td>$11</td>
</tr>
<tr>
<td>Video</td>
<td>$80</td>
</tr>
<tr>
<td>Video 1</td>
<td>$121</td>
</tr>
<tr>
<td>Video 2</td>
<td>$35</td>
</tr>
<tr>
<td>Video 3</td>
<td>$213</td>
</tr>
<tr>
<td>Video 4</td>
<td>$23</td>
</tr>
<tr>
<td>Video 5</td>
<td>$8</td>
</tr>
<tr>
<td>Video 6</td>
<td>$4</td>
</tr>
<tr>
<td>Video 1 - LG 4</td>
<td>$143</td>
</tr>
<tr>
<td>Video 2 - LG 5</td>
<td>$58</td>
</tr>
<tr>
<td>Video 3 - LG 6</td>
<td>$213</td>
</tr>
<tr>
<td>Video 4 - LG 7</td>
<td>$23</td>
</tr>
<tr>
<td>Video 5 - LG 8</td>
<td>$8</td>
</tr>
<tr>
<td>Video 6 - LG 9</td>
<td>$4</td>
</tr>
<tr>
<td>Video 1 - Agg 1</td>
<td>$105</td>
</tr>
<tr>
<td>Video 2 - Agg 2</td>
<td>$103</td>
</tr>
<tr>
<td>Video 3 - Agg 3</td>
<td>$58</td>
</tr>
<tr>
<td>Video 4 - Agg 4</td>
<td>$53</td>
</tr>
<tr>
<td>Video 5 - Agg 5</td>
<td>$39</td>
</tr>
<tr>
<td>Video 6 - Agg 6</td>
<td>$126</td>
</tr>
<tr>
<td>Video 1 - Agg 7</td>
<td>$75</td>
</tr>
<tr>
<td>Video 2 - Agg 8</td>
<td>$57</td>
</tr>
<tr>
<td>Video 3 - Agg 9</td>
<td>$39</td>
</tr>
<tr>
<td>Video 4 - Agg 10</td>
<td>$11</td>
</tr>
</tbody>
</table>

**Remarketing**
- Guaranteed: $80
- Non-guaranteed: $30
- Auto / insurance: $18

**Aggregator insurance**
- Agg 1: $105
- Agg 2: $103
- Agg 3: $58
- Agg 4: $53
- Agg 5: $39
- Agg 6: $126
- Agg 7: $25
- Agg 8: $5
- Agg 9: $5
- Agg 10: $4

**Aggregator lead gen**
- LG 1: $190
- LG 2: $163
- LG 3: $87
- LG 4: $74
- LG 5: $32
- LG 6: $29
- LG 7: $11

**Calibration layer**

**National media (TV & radio)**: $140

**Local media (TV & radio)**: $200

**Direct mail**: $180

**Digital**: $83
How it Works

Hierarchical measurement approach

• Top down media mix model defines overall media performance
• Detailed media mix model just within digital drills down to program
• A third layer could be introduced if need to report down to even more granular level (like creative testing)

Model built at the zip (or IP block) by day-level data

• IP to zip matching is not always accurate, but IP blocks are consistent which allows us to get around cookie deletion problems
• We found that over 90% of cookies only had IP addresses within a single IP block

Created new media mix model approach at very granular level

• Zip/IP block by day level is far more granular than typical media mix models
• The granularity allowed us to:
  - Estimate each of the 70+ digital programs directly in the model
  - Can use only 1 month of data to estimate model – very important to use shorter time frame in quickly changing media like digital media
Modeling Framework Review

**Data:**
- Top Down: All media; multiple sources
- Bottom Up: Digital only; DFA impressions & clicks

**Response Measure:**
- Top Down: Total Inquiries
- Bottom Up: Internet Inquiries

**Timeframe:**
- Top Down: 2 years ending June 20th, 2011
- Bottom Up: Latest 6-months

**Frequency:**
- Top Down: Weekly
- Bottom Up: Daily

**Geography:**
- Top Down: DMA (210)
- Bottom Up: IP-Groups (14,737)

**Methodology:**
- Top Down: Ridge Regression
- Bottom Up: Ridge Regression

**Time Effects:**
- Captured via exponential decay

**Diminishing Returns:**
- Captured via power terms

**Validation:** We tested over 75 different bottom-up model permutations to test estimate sensitivity.
Key Insights

<table>
<thead>
<tr>
<th>Top of funnel media (display/video) produces better CPI than bottom of funnel tactics (search/lead gen)</th>
<th>Results were dramatic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• We estimated a 50%+ error rate for some last click based program estimates.</td>
</tr>
<tr>
<td></td>
<td>• Bottom of funnel programs (such as search and lead gen) metrics were most likely to be overstated.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>One specific category of display showed the lowest return on investment on a CPI basis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Several targeted display programs and some lead gen programs were among the highest performing.</td>
</tr>
<tr>
<td></td>
<td>• Some branded search keywords found to be too expensive given ROI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Video was performing much better than anyone expected on a CPI basis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Some video programs were performing more than 4x better than previously understood based on click rates.</td>
</tr>
<tr>
<td></td>
<td>• However, overall large spread in performance depending on video tactic</td>
</tr>
</tbody>
</table>
Overall Assessment

- Net result: Approach has many advantages but result interpretation is not as intuitive for marketers as with pure cookie modeling based approaches.

Approach Strengths

- Yields incremental estimates
- Accounts for online and offline media effects
- Not sensitive to cookie deletion or cookie privacy rules
- Consistency: Aligns top level marketing mix results with bottom level campaign estimates

Approach Weaknesses

- Can feel like a ‘black box’ to marketers
- Not as intuitive as cookie level modeling approach to marketer
- In future reliability of IP blocks may degrade as mobile device usage increases
Summary

• Digital data is messy and measurement is challenging
• Even though perfect measurement isn’t possible, there is huge value from evolving digital measurement within an organization (get closer to the truth)
• Create a roadmap for what you need to measure
  – Needs
  – Priorities
  – Level of effort and impact
• Use a combination of testing and modeling to address measurement needs
Discussion Questions (Time Permitting)

• How would you rate your organization’s digital measurement sophistication?
• What is preventing your organization from having more effective digital measurement?
• What challenges and successes have you or others in your organization had with running tests within digital media?
• What are you taking away from this discussion as a follow up in your company?
Merkle Inc.

Peter Vandre
Vice President, Digital Analytics Practice
Phone Number: 443.542.4345
Email Address: pvandre@merkleinc.com