Topics

- Data inspiration: Test&Target 1:1
  - Introduction to Test&Target 1:1
  - Test
  - Target
- Description of data
- Theory and practice
  - Pascal data vs. actual data
  - Business Constraints
What does Test&Target 1:1 Do?

- Part of Adobe’s online marketing suite

- Records visitor activity on the site, building up a profile of each individual visitor

- Tracks responses to content, both for individuals and the population as a whole
  - Learns what is most popular for the population as a whole, and watches for time variation – Generalized Learning
  - Builds models and automatically learns what products an individual visitor is most likely to be interested in – Targeting

- Learns dynamically with minimum human analysis needed
Visitor Profile Captured in Server

Site Behavior Variables
- Customer/prospect
- New/return visitor
- Previous visit patterns
- Previous Product interests – top level
- Previous Product interests – low level
- Searches
- Previous online purchases
- Previous Campaign exposure
- Previous Campaign responses

Environment Variables
- IP address
- Geodem/Psychographics
- Country of origin
- Time zone
- Operating system
- Browser type
- Screen resolution

Temporal Variables
- Time of day
- Day of week
- Recency
- Frequency

Referrer Variables
- Referring domain
- Campaign ID
- Affiliate
- PPC
- Natural search
- Direct/bookmark
Process Variable Response Rates

- Age
- Depreciated interest
- Time
- Search data
- Connection speed
- Previous exposure
- Mortgage
- ISP
- Referring site
- Language
- Mortgage area
- Online login
- Recency
- Pension
- Savings
- City
- Rates Area
- Repeat visits
- Salary
- Session depth
And Variable Interactions Response Rates
Combine with Generalized Content Response Rates

- Mortgage
- Loan
- Credit card
- Savings

Response rate
Testing Serve Group

- A small percentage of what is served to the user is purely for testing
  - Percentage is pre-set

- Shown randomly selected content
  - Creates baseline performance for modeling and reporting
  - Provides relative response rates independent of user profile
  - Tracks trends in advertising options – leads discovery of previously mentioned generalized response rates
Targeting Serve Group

- Shown targeted content based on values in the user’s profile

- Recipe with the highest predicted likelihood of response is served
  - When serve decisions are statistically insignificant, model elects to explore
    - In exploration, the model will serve an offer to gather performance information for specific user profiles
  - Otherwise, the user profile is leveraged to provide the maximum performance uplift
### Demo 1:1 Test

**Status:** Approved  
**Start/End:** 02/04/10 08:33PM | when deactivated

#### Metrics Filter
- Extreme Order Filter
- Week Day Filter
- Counting methodology: Visits
- Success metric: Conversion
- Evaluate success by: Conversion Rate

#### Reports
- Select date: Last two months
- Time period: remember report settings
- CSV download

#### Summary

<table>
<thead>
<tr>
<th>Visits</th>
<th>Impressions</th>
<th>Conversion</th>
<th>Lift</th>
<th>AOV</th>
<th>RPV</th>
<th>Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Campagne Totals</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>334,556</td>
<td>424,556</td>
<td><strong>11.15%</strong> ↑ <strong>2.75%</strong></td>
<td></td>
<td>$1,178.00</td>
<td>$38.15</td>
<td>$75,119.52</td>
</tr>
<tr>
<td><strong>Targeted Traffic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>89.35% (298,926)</td>
<td>88.12% (374,161)</td>
<td>12.78% (42,689) ↑ 2.75%</td>
<td></td>
<td>$1,156.87</td>
<td>$41.41</td>
<td>$82,341.93</td>
</tr>
<tr>
<td><strong>Testing [Control] Traffic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.66% (35,630)</td>
<td>11.88% (50,437)</td>
<td>10.92% (3,891) --</td>
<td></td>
<td>$1,126.20</td>
<td>$38.03</td>
<td>$23,789</td>
</tr>
</tbody>
</table>
Description of Data

- Click through(0,1) data collected for a single client between June 2009 to April 2010
- Stratified sampling at 0.1 for negative responses to bring up mean positive response rate to 1% for the sake of creating a richer data set
- Similar campaigns were merged to create six campaign groups. These campaigns were then sampled to give equal serves for each of the six groups
- This data is then categorized into template ranges
- responder: var1 between 0.3 and 0.6, var2 between 6 and 8 …
- non-responder: var1 between 0.1 and 0.3, var2 between 1 and 6 …
- Use above templates to randomly generate a data set within each variables template value
Theory and Practice: Pascal Data vs. Actual Data

- **Non-stationary levers**
  - As the popularity of offers change over time, the assumption of stationary levers (options) is invalidated. Common multi armed bandit solutions like UCB1 are questionable.

- **Multiple success metrics**
  - Advertisers require options for various optimization types.
  - Rather than just conversion(0,1) as a success metric, advertisers can optimize on continuous responses like revenue per visit or average order value.

- **Smaller positive response rates**
  - Depending on how advertisers define conversions or which success metric they choose, response rates can fall well below 1%.
Theory and Practice: Business Constraints

- Direct model away from optimal serves
  - Push particular offer to ensure a minimum/maximum number of serves (new product launch)
  - Business partitioning (requested home page representation)
  - Experience exclusions (two campaigns can’t be served side by side)

- Exploration vs. Exploitation
  - Exploration requires convincing (Choosing to serve poorer performing options)

- Effects
  - Cannot guarantee the best option is chosen for each user
  - All of these constraints come with an opportunity cost
In Conclusion

- Exploration and exploitation balance is a daily concern at Test&Target 1:1

- Business and science side find an acceptable common ground between desirable business outcomes and optimization

- Actual data larger/messier than Pascal data
  - Requires processing in real time to capture all available signals
Website optimisation

• For a given visitor $v$, choose content to display on a webpage. 1 out of $N$ options.

• Objective: maximise engagement=clicks

• Input: $(v,1) \ldots (v,N)$ pairs

Output: index of the pair for which a click is most likely
The data

• 20 million anonymised records: (visitor feature, option index, click indicator)
• 120 continuous and nominal features
• 6 options, all with same CTR (0.24%)
• For a given $v$, only one $(v,o)$ pair out of $N$ will be in the data
The task

- **Input**: batch of 6 visitor-option pairs
- **Output**: index of the pair most likely to be associated to a click

- If click, get reward of 1
- Only observe reward for selected pair
- Maximise cumulated reward (=score)
Evaluation

• Submit ClickPredictor.jar

• Phase 1: live leaderboard
  - from 14 Mar to 6 May
  - 500,000 batches (~ 6 weeks)
  - no initial knowledge about the data
  - live leaderboard
  - logs: reward, time, memory at each iteration

• Phase 2: only one submission
  - from 13 May to 1 Jun
  - 2,810,084 batches (~ 34 weeks)
  - phase 1 data has been revealed
Resources

• Sun Grid Engine at UCL
• 100ms per batch
• 4GB per node -> 3.5GB JVM -> 1.75GB
# Phase 1

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Institution</th>
<th>Score</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Olivier Nicol</td>
<td>INRIA, SequeL</td>
<td>217</td>
<td>0</td>
</tr>
<tr>
<td>#2</td>
<td>Christophe Salperwyck</td>
<td>Orange Labs</td>
<td>207</td>
<td>2</td>
</tr>
<tr>
<td>#3</td>
<td>Aurélien Garivier</td>
<td>CNRS / Telecom ParisTech</td>
<td>204</td>
<td>7</td>
</tr>
<tr>
<td>#4</td>
<td>Olivier Cappé</td>
<td>CNRS / Telecom ParisTech</td>
<td>203</td>
<td>1</td>
</tr>
<tr>
<td>#5</td>
<td>Jérémie Mary</td>
<td>INRIA, SequeL</td>
<td>198</td>
<td>7</td>
</tr>
<tr>
<td>#6</td>
<td>Tanguy Urvoy</td>
<td>Orange Labs</td>
<td>171</td>
<td>4</td>
</tr>
<tr>
<td>#7</td>
<td>Martin Antenreiter</td>
<td>MUL</td>
<td>166</td>
<td>9</td>
</tr>
<tr>
<td>#8</td>
<td>Ronald Ortner</td>
<td>MUL</td>
<td>164</td>
<td>4</td>
</tr>
</tbody>
</table>
## Phase 1

<table>
<thead>
<tr>
<th>Rank</th>
<th>Institution</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>INRIA, SequeL</td>
<td>2170</td>
</tr>
<tr>
<td>#2</td>
<td>Orange Labs</td>
<td>2072</td>
</tr>
<tr>
<td>#3</td>
<td>CNRS / Telecom ParisTech</td>
<td>2047</td>
</tr>
<tr>
<td>#4</td>
<td>MUL</td>
<td>1669</td>
</tr>
<tr>
<td></td>
<td>Random</td>
<td>1177</td>
</tr>
</tbody>
</table>
Phase 1

- Random
- CNRS
- Orange
- INRIA
Phase 1

- CNRS
- Orange
- INRIA
Phase 2

#1 INRIA, SequeL 11529
#2 Orange Labs 10419
#3 CNRS / Telecom ParisTech 9990
#4 MUL 8049
Random 5598
Congratulations!
Phase 2

- Random
- CNRS
- Orange Labs
- INRIA

- Phase 2 submission
- Phase 1 submission (baseline)
Uplift

- Uplift = \frac{s}{r} - 1
  where s is the score of the algorithm and r is the score of random

- 0% if not using visitor features

- 106% for the INRIA algorithm
<table>
<thead>
<tr>
<th>Rank</th>
<th>Name</th>
<th>Affiliation</th>
<th>Total time</th>
<th>Score</th>
<th>Uplift</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Olivier Nicol</td>
<td>INRIA</td>
<td>3h 40m</td>
<td>11529</td>
<td>106%</td>
</tr>
<tr>
<td>#2</td>
<td>Christophe Salperwyck</td>
<td>Orange</td>
<td>29h 50m</td>
<td>10419</td>
<td>86%</td>
</tr>
<tr>
<td>#3</td>
<td>Tanguy Urvoy</td>
<td>Orange</td>
<td>4h</td>
<td>10179</td>
<td>82%</td>
</tr>
<tr>
<td>#4</td>
<td>Aurélien Garivier</td>
<td>CNRS</td>
<td>1h 17m</td>
<td>9990</td>
<td>78%</td>
</tr>
<tr>
<td>#5</td>
<td>Martin Antenreiter</td>
<td>MUL</td>
<td>20h</td>
<td>8049</td>
<td>44%</td>
</tr>
<tr>
<td></td>
<td>Random</td>
<td></td>
<td>1h 12m</td>
<td>5598</td>
<td>0%</td>
</tr>
</tbody>
</table>
• Stochastic algorithms
• Batches presented in the same order, but elements in a batch presented in different orders at each evaluation
Resources

• 78h theoretical max running time (100ms per batch)
• the INRIA algorithm only took 3h 40m
• 1.75GB of memory available, 20GB of data