Digital Fingerprinting in Market and Survey Research

CASRO Panel Conference

Duane Berlin

Peter Milla

dberlin@levberlin.com

pmilla@optonline.net
Topic List

• Background:
  • Research Industry Quality Concerns
  • MRIA investigation
• Ponemon/ThreatMetrix Research
• Digital Fingerprinting for Research
• Digital Fingerprinting and Privacy
• CASRO Efforts
• Draft Guideline on the Use of Digital Fingerprinting

Note: Digital Fingerprinting, the term used in survey and market research, is also known as Device Fingerprinting, Device ID or Machine ID.
Background

• **Research Industry Quality Concerns:**
  – Recent quality concerns having to do with online panel/online sample sources:
    • Duplication
    • Fraud/respondent quality
  – Industry associations have mobilized with various efforts/programs
  – Digital Fingerprinting (DF, also referred to as Device Fingerprinting, Device Id or Machine-Id) has emerged as an effective solution to address the duplication and fraud/respondent quality problems

• **MRIA (Market Research and Intelligence Association) in Canada:**
  – Legal opinion on potential PIPEDA violations
  – Driven partially by disclosure issue (lack of adequate disclosure)
Ponemon/ThreatMetrix Research – “Online Consumers Reaction to Device Fingerprinting”, 2009

• Ponemon Institute, LLC:
  – Ponemon Institute is dedicated to independent research and education that advances responsible information and privacy management practices within business and government. Ponemon’s mission is to conduct high quality, empirical studies on critical issues affecting the management and security of sensitive information about people and organizations.

• ThreatMetrix, Inc:
  – ThreatMetrix device identification (www.threatmetrix.com) helps companies control online fraud and abuse in real time so they can acquire more customers faster, reduce costs, and increase customer satisfaction. ThreatMetrix profiles the device used in an online transaction so companies can determine whether the users are fraudsters or customers.
Experience with Online Fraud
(Base = Victims of Online Fraud – 43%)

- Unwanted online ads or email solicitations: 88%
- Fraudulent charges on my credit card: 24%
- Theft of financial assets, including bank account checking/savings account: 6%

Ponemon Institute© Private and Confidential Report
Sponsored by ThreatMetrix
Level of Concern About Online Fraud

- **Very Concerned**: 26%
- **Concerned**: 35%
- **Somewhat Concerned**: 19%
- **Not Concerned**: 20%
What Consumers Want

- Require more identification from consumers: 21%
- Use technology to improve authentication of consumers: 78%
- Increased efforts to prevent fraudsters from stealing consumers' information: 83%
Consumers Favorable Impression about Device Fingerprinting (Base = Would Agree to Device Fingerprinting – 69%)

- It would be more convenient because I would not have to remember passwords or answered pre-selected questions (75%)
- It would increase security and protection of my personal information (64%)
- I would have to provide less personal information (19%)

Ponemon Institute© Private and Confidential Report Sponsored by ThreatMetrix
### Data Types Consumers are Willing to Share in Order to Authenticate Device

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial number of computer</td>
<td>91%</td>
</tr>
<tr>
<td>Type and make of computer</td>
<td>82%</td>
</tr>
<tr>
<td>Internet service provider</td>
<td>80%</td>
</tr>
<tr>
<td>Browser settings</td>
<td>72%</td>
</tr>
<tr>
<td>Type of browser</td>
<td>69%</td>
</tr>
<tr>
<td>IP address</td>
<td>63%</td>
</tr>
<tr>
<td>Types of software on device</td>
<td>52%</td>
</tr>
</tbody>
</table>
Digital Fingerprinting for Research

• What is Digital Fingerprinting?
  – Technologies that deploy an algorithm that analyzes a large number of technical characteristics and settings to generate a unique identifier that can identify a specific computer (a Machine-ID or Device Id)
  – Algorithms typically consist/utilize many variables, which can include:
    • IP address
    • O/S type/version
    • Browser type/version
    • Browser plug-ins
    • Browser parameter values
    • Cookies
    • Flash local objects
  – Algorithms generally perform very will and are reliable
  – Algorithm components are not PII (more later)
  – Many solutions on the market, all with high accuracy (many cite > 99%)
Digital Fingerprinting for Research (continued)

• Online banking use:
  – In addition to username, password, etc, a digital fingerprint is used by many banks to identify the device being used as verification as the assumption that most people use the same computer for banking.
  – The technology can be used to permit legitimate customers in and also block “bad actors.”
  – Digital fingerprints can/are stored and can be utilized on subsequent sessions.
Digital Fingerprinting Solutions for Research
Include

• Bloodhound from Western Wats
• MKTG Crop Duster
• mo’web Research
• Peanut Labs OptimusID
• RelevantView RelevantID
• TrueSample via 41st Parameter (MarketTools)
• Proprietary solutions
Digital Fingerprinting – Related Technologies

• Some research companies incorporate Digital Fingerprinting technology with other technologies/technology approaches or use other technologies in a singular or hybrid fashion

• Other technologies include:
  – Email encryption (email hashing)
  – Post-survey data verification
  – Other (IP Address validation, cookie approaches)
  – Mixed mode (many companies apply)
Digital Fingerprinting in Research -- Applications

- De-duplication
- Geo-IP fencing/validation
- Proxy server identification
- Survey quality measures (used in conjunction with duplication features), including dealing with fraudulent respondents:
  - Cheating probability scores/indexes
  - Question response patterns
  - Question completion times
  - Challenge response with CAPTCHA
  - Post survey detection (incentive payment, client feedback)

Note: While survey quality measures are typically used in conjunction with Digital Fingerprinting, not all Digital Fingerprinting software provides this functionality, which may be provided by other software.
Digital Fingerprinting De-Duplication

• De-duplication:
  – Within and across studies
  – Within individual source or multiple sources (panels, web intercepts, social networks, etc.)

• Key requirements:
  – Must be able to screen all sample sources
  – Must be compatible with all platforms
  – Must be real time
  – Must not require downloaded software
  – Must not require transfer of PII (Personally Identifiable Information)
Digital Fingerprinting De-Duplication (continued)

• Duplication problem is generally small (< 1% for general population survey for well-managed panels or respondent sources):
  – Issues are manifested with hard-to-reach populations or poorly managed panels or respondent sources
  – Other causes of error are more of the issue
Digital Fingerprinting in Research and Privacy

• Digital Fingerprinting potentially involves privacy issues (thinking to this point that the technology doesn’t/won’t collect PII)

• In some countries outside the U.S., such as EU member nations and Canada (currently under investigation), a computer user’s IP address *may* be considered personal data

• Unique DFs could be linked to identifiable individuals

• Potentially raises data accuracy/data retention (keeping PII accurate and using it only as long as absolutely necessary are two key privacy principles (data changes and data removal))
Digital Fingerprinting in Research and Privacy (continued)

• Machine-IDs (Digital Fingerprint or Device ID) could be compiled with other more sensitive information to create profiles and could be cross-referenced to learn more about panelists/respondents:
  – The research transaction chain could cause issues (panel owner transfers to research company who transfers to end client/other party who may try to scrape identifying information)

• Information could be sent to other parties (data match/append requests)

• Browser string checking/matching could present privacy problems
Digital Fingerprinting in Research and Privacy (continued)

- Practice of research firms/panel companies sending panelists to other research companies/panel companies could cause privacy policy issues if other research companies/panel companies use Digital Fingerprinting.
- Very little privacy protection for this kind of activity, especially in the US.
- Currently seeing activity/significant concern about behavioral advertising tracking which could impact regulators interpretation.
Digital Fingerprinting and Privacy – Recommendations

*Key is to establish practices that protect respondents, practitioners, clients and the industry.*

- Address in privacy policy (also consider notification at survey start or other points where Machine-ID is generated)
- Include privacy framework requirements in research agreements, MSAs, subcontractor agreements, etc.
- Store only essential information for only as long as necessary
- Privacy training for staff, including how data can be used, shared and who can access.
Digital Fingerprinting and Privacy – Recommendations (continued)

• Include in Member/Non-Member Privacy Policy:
  – Identification of the types of information used to create the Machine-ID
  – Information on how Machine-ID will be used (i.e. to identify duplicate respondents/panelists, to flag or block respondents/panelists, etc.)
  – Whether and to what extent these technologies will store and use the Machine-IDs and other information collected from a particular study in a database and/or in other research projects
  – A statement that the company will use commercially-reasonable efforts to ensure that the digital fingerprinting technologies are safe, secure and do not cause undue privacy or data security risks
Digital Fingerprinting and Privacy – Recommendations (continued)

• Include in Member/Non-Member Privacy Policy (continued):
  – A statement that the company will manage the use and distribution of
digital fingerprinting technologies to ensure that the use of such
technologies is conducted in a professional and ethical manner, in
accordance with the CASRO Codes of Standards, etc., sample provider
privacy policies, applicable laws, and statements and disclosures made
to data subjects.
  – A statement that if the company discovers unethical conduct or that
the digital fingerprinting technologies are not being used consistent
with statements and disclosures made to data subjects and applicable
laws, the company will take immediate action to correct, rectify, and
prohibit such unethical conduct and to ensure the proper
administration of such technologies.
CASRO Position/Efforts

• CASRO Position:
  – Digital Fingerprinting (DF, Device Fingerprinting, Device ID or Machine-ID) technology is used by survey researchers as an effective quality control that maintains the integrity of web-based research.
  – Like any other computing technology, Digital Fingerprinting (DF) must be employed responsibly and transparently -- consistent with personal and data privacy laws and in accordance with ethical and professional standards.
  – The use of Digital Fingerprinting pursuant to standards and guidelines that appropriately protect respondent privacy rights is an ethical practice.
  – Such use of DF technology is consistent with US privacy and data protection laws.
CASRO Position/Efforts (continued)

- CASRO Position (continued):
  - More data, research and input from providers and users of DF technology is needed before it can be determined whether the use of such technology complies with the privacy regulations in other jurisdictions.
  - Disposition towards this technology is to analyze whether the technology can be properly and ethically used in research, and, if so, to establish standards and guidelines for the use of the technology.
  - CASRO believes online research data quality is a critical issue and a major focus for research businesses, clients, and associations. The research industry, and the associations that serve it, must consider and evaluate the efficacy of all technologies and solutions that support the quality of online research.
  - Potential coming use of panelist/respondent validation will need to be part of this discussion.
CASRO Efforts (continued)

• CASRO Efforts:
  – Investigating technological issues (including examining all alternative technologies, their efficacy and privacy implications)
  – Working with research providers, research service providers and technology vendors on:
    • Best practices for disclosure
    • Draft Guideline
Draft Guideline on the Use of Digital Fingerprinting

- Companies involved to date:
  - e-Rewards/Research Now
  - Burke
  - GfK
  - GMI
  - Imperium (RelevantID)
  - Harris Interactive
  - IPSOS
  - Lightspeed (Kantar)
  - MarketTools (TrueSample)
  - MKTG (Crop Duster)
  - MKTG (Crop Duster)
  - NPD
  - Peanut Labs (Optimus)
  - Ponemon Institute
  - SSI
  - Synovate
  - Toluna/Greenfield Online)
  - Western Wats (Bloodhound)

[Others are welcome]
Draft Guideline on the Use of Digital Fingerprinting (continued)

• Initial review of current draft by participating companies on 9 Feb.:  
  – Disclosure issue covered in depth, preferred path is disclosure via privacy policy  
  – Focus on fraud prevention  
  – Not about prohibiting technology, but managing its use  
  – Self-regulation a key theme

• Current draft includes:  
  – Definition  
  – Digital fingerprinting and survey respondent’s privacy  
  – Proper use of digital fingerprinting mean  
  – Survey Participation Databases that use digital fingerprinting or other respondent identification technologies
Draft Guideline Review

- Review of current draft
- Next steps
Digital Fingerprinting in Market and Survey Research

CASRO Panel Conference
Duane Berlin
Peter Milla

dberlin@levberlin.com
pmilla@optonline.net