Phishing Defense against IDN Address Spoofing Attacks

Viktor Krammer ¹,²
1 E-Commerce Competence Center
2 Vienna University of Technology

http://www.quero.at/
Qui quaerit, invenit.

Biblia Vulgata, Lc 11, 9
Agenda

- About Phishing
- Internationalized Domain Names (IDN)
- Address Spoofing Attacks
- IDN-based Attacks
- Anti-Spoofing Techniques
- Conclusion
How we authenticate ...

People
- Name (first, last, middle, nickname)
- Appearance (face, eyes, body, clothes)
- Voice, Gestures, Behavior
- Documents (driving licence, passport, ID cards)
How we authenticate ...

Web sites
- Address (URL, host name, domain name)
- Appearance (page layout, design, logo, colors)
- Content
- Certificate (SSL/TSL)

Addresses per se unique, but not for humans!
- confusingly similar
- likelihood of confusion
Phishing Example (1)

Target: Bank Austria

Sender:

Real Address: online.ba-ca.com

Fake Address: ba-ca.onlinebanking.com.de.ronounced.tk
Phishing Example (2)

Real

Quero
Why Phishing Works

User study conducted by Rachna Dhamija et al. (presented at CHI 2006)

- 22 participants classifying 20 Web sites

Key Findings

- 23% of participants looked only at the content to authenticate the Web site
  → 77% looked at the address bar or at other security indicators
- 90% were fooled by well-crafted Phishing sites
Agenda

✓ About Phishing
  ▪ Internationalized Domain Names (IDN)
  ▪ Address Spoofing Attacks
  ▪ IDN-based Attacks
  ▪ Anti-Spoofing Techniques
  ▪ Conclusion
Internationalized Domain Names

- Defined in RFC 3490, 3491, 3492, 3454, 3743, 3987
- Client-side extension of DNS allowing non-ASCII characters in domain names
- Based on a subset of Unicode 3.2
- Backward compatible with existing DNS (IDN labels are stored in ASCII beginning with “xn--”)
- Uses Punycode for encoding Unicode domain labels efficiently (run-length compression)

Example
東京理科大学.jp
Punycode xn--1lq68wkwbj6ugkpiji.jp
UTF-8 %E6%9D%B1%E4%BA%AC%E7%90%86%E7%A7%91%E5%A4%A7%E5%AD%A6.jp
Brief history of IDN (1)

- 2002 The Homograph Attack (Gabrilovich, Gontmakher)
- 2002 Unicode 3.2
- 2003 VeriSign releases i-Nav plug-in for IE5/6
- 2003 ICANN publishes IDNA RFCs
- 2003 Opera 7.11 adds IDN support
- 2004 Mozilla adds IDN support to Firefox 0.8
- 2005 February IDN security receives big media coverage resulting from an article by Shmoo Group exploit: “own any domain, no defense exists”
- 2005 July Unicode Security Considerations rev.3
Brief history of IDN (2)

- **2005 July**
  Quero Toolbar 2.1 released with IDN script highlighting and mixed-script security warnings

- **2005 November**
  Quero Toolbar 2.2 reached RFC-compliance

- **2006 October**
  Microsoft released IE7 with native IDN support, mixed-script detection and an integrated Phishing filter
Agenda

- About Phishing
- Internationalized Domain Names (IDN)
  - Address Spoofing Attacks
  - IDN-based Attacks
  - Anti-Spoofing Techniques
  - Conclusion
Address Spoofing Attacks (1)

User Confusion-based Attacks

- Confusion by Name Similarity
  Example: southtrustonlines.com vs. southtrust.com

- Confusion by Address Complexity
User Confusion-based Attacks (cont’d)

- Confusion by Random Addresses
  Example:
  http://secure-user-survey.com/exec/obidos/subst/home/sv/

Vulnerability-based Attacks

- Client-side Vulnerabilities
- Server-side Vulnerabilities
Agenda

✓ About Phishing
✓ Internationalized Domain Names (IDN)
✓ Address Spoofing Attacks
  • IDN-based Attacks
  • Anti-Spoofing Techniques
  • Conclusion
### IDN-based Attacks (1)

- **Mixed-script Spoofing**
  - Substituting characters with visually similar ones

<table>
<thead>
<tr>
<th>Latin</th>
<th>Cyrillic</th>
<th>Greek</th>
</tr>
</thead>
<tbody>
<tr>
<td>P 0050</td>
<td>P FF30</td>
<td>P 03A1</td>
</tr>
<tr>
<td>p 0070</td>
<td>p FF50</td>
<td>ρ 03C1</td>
</tr>
<tr>
<td>S 0053</td>
<td>S FF33</td>
<td>Σ 03A3</td>
</tr>
<tr>
<td>s 0073</td>
<td>s FF53</td>
<td>ζ 03C2</td>
</tr>
<tr>
<td>T 0054</td>
<td>T FF34</td>
<td>T 03A4</td>
</tr>
<tr>
<td>t 0074</td>
<td>t FF54</td>
<td>τ 03C4</td>
</tr>
</tbody>
</table>
Whole-script spoofing
using characters from only one script that are reinterpreted in another script
Example: Latin caxap.ru
        Cyrillic caxap.ru (xn--80aa2cbv.ru)

Single-script spoofing
exploiting similarities between characters within one script
0 vs. o; l vs. t; 1 vs. l; m vs. rn; etc.
IDN-based Attacks (3)

- Syntax Spoofing
  \/ (U+002F) vs. /(U+2044), / (U+2215)

- Numeric Spoofing
  8 (U+0038) vs. U+09E6, U+09EA

- Invisible Character Injection
  Control, formatting, tagging and spacing characters are prohibited by IDN Nameprep

- Bidirectional Text Spoofing
  Eliminated by IDN Nameprep
IDN-based Attacks (4)

- **Combining Mark Order Spoofing**
  encoding specific threat: order of combining marks can be ambiguous

- **Inadequate Rendering Support**
  Example: repeating combining marks
  `<c, a, f, e, U+0301, U+0301>` looks like
  `<café>`
  ṭ U+301 Combining Acute Accent
Agenda

✓ About Phishing
✓ Internationalized Domain Names (IDN)
✓ Address Spoofing Attacks
✓ IDN-based Attacks
  ▪ Anti-Spoofing Techniques
  ▪ Conclusion
Anti-Spoofing Requirements

- RFC Compliance
- Avoiding Discrimination
- Preferring Self-contained Solutions
- Alerts
- Appropriate Rendering Support
- User Preferences (allow opt-out, Whitelist)
Visualisation Techniques (1)

- **Digit Indication**

- **IDN Indication & Highlighting**
  - Characters from different script groups receive different background colors
  - Displaying the names of the script groups to mitigate whole-script attacks
Visualisation Techniques (2)

- Secure Connection Indication
- Core Domain Highlighting

„Core Domain“: most relevant part of the address usually: 2nd and 1st level domain label
UI Improvements

- Address Bar Integration
- Security Related Information:
  - Current Location (URL)
  - Core Domain
  - Secure Connection Icon (Certificate Details)
  - Blocked Content
  - Security Warnings
- Support for Larger Font Sizes (default: 8 pt!)
- Switching to ACE Form
Security Warnings

- **Invalid Addresses**
  not well-formed according to RFC definition

- **Suspicious Character Detection**
  alerts the user in cases of mixed-script

  **Assumptions:**
  - harder to exploit similarities within one script
  - rather undesirable to mix scripts (harder to input, read, recognize and memorize)

- **Missing Glyph Detection**
Quero Toolbar

- Add-on for Internet Explorer
- RFC-compliant implementation of IDN standards
- Adds IDN support to older versions of IE
- Demonstrating anti-spoofing techniques
- New user interface (combines search and navigation into one toolbar)
- Integrated content filter
- Over 10,000 times downloaded (2005/01–2006/01)
- Freeware licence
Agenda

✓ About Phishing
✓ Internationalized Domain Names (IDN)
✓ Address Spoofing Attacks
✓ IDN-based Attacks
✓ Anti-Spoofing Techniques

- Conclusion
Conclusion

- Besides the padlock icon the address is still the most important indicator for authenticating a Web site.
- Spoofed addresses are no longer visually distinguishable from their legitimate counterpart.
- Quero helps the expert and non-expert user to make better trust decisions based on the current URL.
- Major Web browser vendors have adopted mixed-script detection and included a blacklist-based phishing filter.
Viktor Krammer
feedback@quero.at
http://www.quero.at/