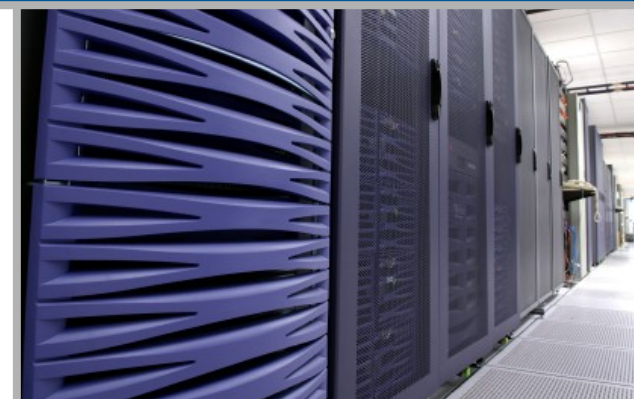




Trustwave[®]

Information Security & Compliance



Advanced Techniques in Automated Web Application Testing

David Byrne
Application Penetration Testing
dbyrne@trustwave.com

Importance of Browser Emulation

- Browser-targeted attacks (XSS)
- HTML component analysis
- Spidering

Old-Style Test

```
http://www.example.com/index.php?  
input="></SCRIPT></FORM><SCRIPT>alert(document.c  
ookie)</SCRIPT><SCRIPT><FORM>
```

- Attempts to escape from several HTML output contexts
- Sends simple XSS attack payload

HTML Output Contexts

- HTML text node
- HTML tag attribute name
- HTML tag attribute value
- HTML event handler value
- HTML comment
- HTML tag name
- TITLE tag block
- PRE tag block
- TEXTAREA tag block
- FORM tag block
- SCRIPT tag block
- STYLE tag block
- IFRAME tag block
- Etc.

Identifying Output Contexts

...

```
<INPUT name="blah" value="zzz">
```

...

Identifying Output Contexts

...

<!--

...

<INPUT name="blah" value="ZZZ">

...

-->

...

Identifying Output Contexts

...

```
<INPUT name="confusing" value="<!--">
```

```
<INPUT name="blah" value="ZZZ">
```

```
<!--
```

```
ha ha ha
```

```
-->
```

...

Identifying Output Contexts

...

```
<textarea>
```

```
<INPUT name="confusing" value="<!--">
```

```
<INPUT name="blah" value="ZZZ">
```

```
<!--
```

```
ha ha ha
```

```
-->
```

```
<textarea>
```

...

Old-Style Test

```
http://www.example.com/index.php?  
input="></SCRIPT></FORM><SCRIPT>alert(document.cookie)</SCRIPT><SCRIPT><FORM>
```

- Simple text matching will result in false positives for many output contexts.
- Doesn't account for different quote types
- Has no intelligent evasion techniques

HTML Document Object Model

- Standardized (w3.org/DOM) object-oriented API for HTML documents and behavior
- Present in all full-featured browsers
- Used by the browser to render the server's response into viewable content
- Used by JavaScript for interactive webpage content
- Grendel
 - HTML parsing engine & DOM implementation: Cobra (Lobo Browser)
 - JavaScript Engine: Mozilla Rhino (Java port of Firefox engine)

Advanced XSS Testing

- **Seed inputs with tokens; identify input/output flows**
- Identify output context
- Test flow for filtered characters
- Identify quote type, if relevant
- Craft and send attacks
- Run executable content in responses

Advanced XSS Testing

- Seed inputs with tokens; identify input/output flows
- **Identify output context**
 - DOM is tree-based
 - Nodes contain type, name, value, child nodes
 - Tree-search simple to implement
- 3. Test flow for filtered characters
- 4. Identify quote type, if relevant
- 5. Craft and send attacks
- 6. Run executable content in responses

Advanced XSS Testing

- Seed inputs with tokens; identify input/output flows
- Identify output context
- **Test flow for filtered characters**
 - < > ' " etc
- 4. Identify quote type, if relevant
- 5. Craft and send attacks
- 6. Run executable content in responses

Advanced XSS Testing

- Seed inputs with tokens; identify input/output flows
 - Identify output context
 - Test flow for filtered characters
 - **Identify quote type, if relevant**
 - Single, double or none
5. Craft and send attacks
 6. Run executable content in responses

Advanced XSS Testing

- Seed inputs with tokens; identify input/output flows
- Identify output context
- Test flow for filtered characters
- Identify quote type, if relevant
- **Craft and send attacks**
- Run executable content in responses

Grendel's XSS Testing

- Seed inputs with tokens; identify input/output flows
- Identify output context
- Test flow for filtered characters
- Identify quote type, if relevant
- Craft and send attacks
- **Run executable content in responses**
 - SCRIPT blocks
 - References to external script files
 - HTML event handlers: OnClick, OnLoad, OnError, etc
 - Custom DOM extension

Response Comparison

- Logical file-not-found detection
- SQL tautologies
- Directory traversal
- Logged-out detection

Levenshtein Distance

- Vladimir Levenshtein, 1965
- Number of character changes (insertions, deletions, substitutions) to transform one string into another
- "TAXI" -> "CAB" = 3

Levenshtein Distance

- Vladimir Levenshtein, 1965
- Number of character changes (deletions, additions, substitutions) to transform one string into another
- "TAXI" -> "CAB" = 3

TAXI

Levenshtein Distance

- Vladimir Levenshtein, 1965
- Number of character changes (deletions, additions, substitutions) to transform one string into another
- "TAXI" -> "CAB" = 3

CAXI

Levenshtein Distance

- Vladimir Levenshtein, 1965
- Number of character changes (deletions, additions, substitutions) to transform one string into another
- "TAXI" -> "CAB" = 3

CABI

Levenshtein Distance

- Vladimir Levenshtein, 1965
- Number of character changes (deletions, additions, substitutions) to transform one string into another
- "TAXI" -> "CAB" = 3

CAB

Grendel's Comparison Techniques

- Tracks actual score & maximum possible score
- HTTP response code – 150
- MIME type – 150
- Set-Cookie name – 100
- Skewed Levenshtein distance of HTTP location headers - 100
- Levenshtein distance of normalized HTML text nodes (sloowwww) - 100
- HTML tag count ratios (min count / max count)
 - APPLET - 50
 - OBJECT - 50
 - EMBED - 50
 - TABLE - 20
 - TR - 20
 - A - 10
 - LINK - 10
 - IMG - 10

Other Possible Comparison Metrics

- Content length
- Human language word count (total and per word) of HTML text nodes, may be faster than a Levenshtein distance
- Character frequency

Logical File-Not-Found Messages

- Request random file names with common platform extensions (.aspx, .jsp, .php, etc).
- Compare response to known FNF messages; add if new
- Compare all responses to known profiles; identify as FNF if similarity exceeds predefined threshold

www.grendel-scan.com