

ThreatFactor NSIA - Feature #227

Ability for ThreatScripts to Add to URL Scan List

10/25/2010 10:42 AM - Luke Murphey

Status:	Closed	Start date:	11/05/2010
Priority:	Normal	Due date:	11/05/2010
Assignee:	Luke Murphey	% Done:	100%
Category:	Scan Engine	Estimated time:	3.00 hours
Target version:	1.0.1		
Description			
<p>ThreatScripts ought to be able to add URLs to the list to scan. This is a nice feature because this allows the resource extraction to be modified with updated definitions which could extract links from the robots.txt files, PDFs, CSS files, etc.</p> <p>Disabling the scripts would thus disable extraction of the relevant links. Below would be an example of the code snippet:</p> <pre>var url = "http://google.com"; ScanURLs.add(url);</pre>			
Related issues:			
Blocks ThreatFactor NSIA - Feature #62: Parse CSS and JavaScript in Detection...		New	04/08/2010

History

#1 - 10/25/2010 11:06 AM - Luke Murphey

- Target version changed from 1.0 (Release) to 1.0.1

#2 - 10/26/2010 01:00 AM - Luke Murphey

- Assignee set to Luke Murphey

#3 - 10/26/2010 01:21 AM - Luke Murphey

- Due date set to 11/05/2010

- Start date changed from 10/25/2010 to 11/05/2010

- Estimated time set to 3.00 h

#4 - 11/01/2010 11:53 PM - Luke Murphey

- Category set to Scan Engine

#5 - 11/03/2010 01:49 AM - Luke Murphey

- Status changed from New to In Progress

- % Done changed from 0 to 50

Implemented in r988. Need additional testing to confirm that the loaded URLs are processed by the scan engine.

#6 - 11/03/2010 05:03 PM - Luke Murphey

Initial testing shows that the link extraction definitions do work. Below was the definition tested:

```
/*
 * Name: ScannerSupport.LinkExtraction.Test
 * ID: 1000000
 * Version: 1
 * Message: This is a test
 * Severity: Medium
 */
```

```
importPackage(Packages.ThreatScript);
```

```
importPackage(Packages.HTTP);
function analyze( httpResponse, operation, environment ){
    a = new Array();
    a[0] = new URL("http://Threatfactor.com/TEST");
    return new Result( false, "Definition did not match the input", a);
}
```

Links extracted this way will be added to the scan list even if they do not match the domain name restriction.

#7 - 11/03/2010 05:48 PM - Luke Murphey

- % Done changed from 50 to 70

Implemented methods that allow URLs to be designated as needing to match the domain limit or not in r990.

#8 - 11/03/2010 11:03 PM - Luke Murphey

The changes were tested with the following definition which only tries to access "http://Threatfactor.com/TEST" if the domain does not match but always tries to access "http://Threatfactor.com/TEST_ALL".

```
/*
 * Name: ScannerSupport.LinkExtraction.Test
 * ID: 1000339
 * Version: 1
 * Message: This is a test
 * Severity: Medium
 */

importPackage(Packages.ThreatScript);
importPackage(Packages.HTTP);

function analyze( httpResponse, operation, environment ){
    a = new Array();
    a[0] = new URL("http://Threatfactor.com/TEST");
    result = new Result( false, "Definition did not match the input", a);
    result.addURL( new URL("http://Threatfactor.com/TEST_ALL"), true);
    return result;
}
```

#9 - 11/04/2010 12:33 AM - Luke Murphey

- *Status changed from In Progress to Closed*

- *% Done changed from 70 to 100*

Changed definition such that URLs are only extracted from definitions if they are not filtered out by the scan policy. This feature has been fully implemented in r994.