An Analysis of Address Space Layout Randomization on Windows Vista

Ollie Whitehouse, Architect, Symantec Advanced Threat Research

Abstract: Address Space Layout Randomization (ASLR) is a prophylactic security technology aimed at reducing the effectiveness of exploit attempts. With the advent of Microsoft Windows Vista, ASLR has been integrated into the default configuration of a Microsoft Windows operating system for the first time. We measure the behavior of the ASLR implementation in the Vista RTM release. Our analysis of the results uncovers predictability in the implementation that reduces its effectiveness.

Index Terms: Address Space Layout Randomization, ASLR, Security, Microsoft, Windows, Vista, Visual Studio, Dynamicbase

I. INTRODUCTION

Address space layout randomization, or ASLR, is a prophylactic security technology that increases system security by increasing the diversity of attack targets [20]. Rather than increasing security by removing vulnerabilities from the system, ASLR makes it more difficult to exploit existing vulnerabilities. This technology is complementary to efforts to remove security vulnerabilities since it can offer some protection for vulnerabilities that are not yet known or have not yet been remedied. ASLR is also complementary to other prophylactic techniques such as Data Execution Prevention (DEP) since the combination of both technologies provides a much stronger defense against memory manipulation vulnerabilities than either one alone.

Techniques for exploiting memory manipulation vulnerabilities are sensitive to the memory layout of the program being targeted. This important class of vulnerability includes stack and heap overflows, underflows, format string vulnerabilities, array index overflows and uninitialized variables. By randomizing the memory layout of an executing program, ASLR decreases the predictability of the memory layout and reduces the probability that an individual exploit attempt will succeed. The security offered by ASLR is based on several factors [12], including how predictable the random memory layout of a program is, how tolerant an exploit technique is to variations in memory layout, and how many exploitation attempts an attacker can practically make.

Address space layout randomization has been integrated into and available for several popular operating systems such as OpenBSD and Linux for several years. Third-party ASLR implementations have been available for previous versions of the Microsoft Windows operating systems as stand-alone products or as part of a Host intrusion protection (HIPS) solutions [1][2][4][11][18]. With the advent of Microsoft Windows Vista, ASLR has been integrated into the default configuration of a Microsoft Windows operating system for the first time [5].

We measured the behavior of the address space layout randomization feature in the 32-bit Vista RTM release. This paper presents our measurements and discusses our measurement techniques. Our analysis uncovers some flaws that reduce the effectiveness of Vista’s ASLR implementation.

A. Previous Work

There have been several efforts to describe and measure the ASLR implementation in Vista. The first is described in the tuxedo-es.org blog [13] and supported by the release of their VistaProbe tool [14][15][16][17]. This was followed shortly by a paper by Rahbar [10] but it’s analysis was refuted by Howard [6]. Both efforts used a beta release of Vista to make their measurements since a release version was not yet available.

To our knowledge our work is the first to measure the ASLR implementation in the Vista RTM release. Unlike previous work, we rebooted the test system between measurements to measure ASLR in the environment it was intended to be used in. While previous work relied on a small number of measurements to draw conclusions, we took a much larger number of measurements and expect our results to have more statistical significance. Finally, we report several important and unexpected deficiencies in the ASLR implementation that were not previously reported and have since been acknowledged by Microsoft.

B. Outline

The remainder of this paper is organized as follows: Section II describes the ASLR implementation provided by
Vista and describes our methodology for measuring its behavior. Section III presents and analyzes the measurements we made. Section IV presents our conclusions based on the analysis.

II. ASLR in Vista

Windows Vista provides address space layout randomization on a per-image basis. Any executable image which contains a PE header, such as executable binaries (.exe) and dynamic link libraries (.dll), can elect to participate in address space layout randomization. This election is made by setting a bit (0x40) in one of the PE header fields (DLLCHARACTERISTICS) [7]. An option (/dynamicbase) is provided by the Microsoft Visual Studio 2005 linker for setting this bit when linking an image.

While loading an image that has elected to participate in ASLR the system uses a random global image offset. This offset is selected once per reboot, although we’ve uncovered at least one other way to cause this offset to be reset without a reboot (see Appendix II). The image offset is selected from a range of 256 values and is 64kB aligned. This offset and the other random parameters are generated pseudo-randomly [3]. All images loaded together into a process, including the main executable and DLLs, are loaded one after another at this offset. Because image offsets are constant across all processes a DLL that is shared between processes can be loaded at the same address in all processes for efficiency.

When executing a program whose image has been marked for ASLR the memory layout of the process is further randomized by randomly placing the thread stack and the process heaps. The stack address is selected first. The stack region is selected from a range of 32 possible locations, each separated by 64kB or 256kB (depending on the STACK_SIZE setting). Once the stack has been placed, it the initial stack pointer is further randomized by decrementing it a random amount. The initial offset is selected to be up to half a page (2048 bytes) but is limited to naturally aligned addresses (four byte alignment on IA32 and 16 byte alignment on IA64). The choices result in an initial stack pointer chosen from one of 16384 possible values on an IA32 system.

After the stack address has been selected, the process heaps are selected. Each heap is allocated from a range of 32 different locations, each separated by 64kB. The location of the first heap must be chosen to avoid the previously placed stack, and each of the following heaps must be allocated to avoid those allocated before it.

The address of an operating system structure known as the Process Environment Block (PEB) is also selected randomly. The PEB randomization feature was introduced earlier in Windows XP SP2 and Windows 2003 SP1 and is also present in Windows Vista. Although implemented separately it is also a form of address space randomization but unlike the other ASLR features, PEB randomization occurs whether or not the executable being loaded elected to use the ASLR feature.

An important result of Vista’s ASLR design is that some address space layout parameters such as PEB, stack and heap locations are selected once per program execution. Other parameters, such as the location of the program code, data segment, BSS segment and libraries, change only between reboots.

A. Methodology

We measured Vista’s ASLR implementation to verify its behavior and to determine how random the memory layout of loaded programs is. We constructed a program to log several important addresses associated with the program each time it is executed. To measure the randomization of the image base address the test program prints the address of a function in the code segment. To measure the randomization of the stack, it prints the address of an automatic variable. To measure the randomization of the PEB structure it prints the address of the PEB. Finally it measures the placement of three heaps. It measures the CRT heap by printing the first value returned by malloc. It prints the first values returned by HeapAlloc using the default process heap, and using a heap created with HeapCreate.

We compiled our test program with Microsoft Visual Studio 2005 SP1 Beta and linked it statically. The program source is listed in Appendix I. When executed repeatedly with the same environment on a machine that does not support ASLR or PEB randomization, this program reports the same constants each time. Any variation in the output is due solely to the effects of ASLR.

Some ASLR parameters are only set once per system boot1. To properly measure the effects of ASLR we decided to reboot the system between measurements. We configured an AMD Athlon 3200 system running 32-bit Windows Vista RTM to automatically log in and run our data collection utility during system startup and then to reboot. This setup closely mimics the environment of long running servers which are executed once during system startup. Our test harness was used to collect samples from 11,500 test runs over the course of twelve days. The complete data is being made available [19].

III. Analysis

We analyzed the results of our measurements to quantify the amount of randomness introduced into the memory layout of a process by the ASLR implementation. Figure 1 show the collected data for HeapAlloc addresses plotted as a series of samples. The plot reveals no noticeable patterns, indicating some amount of randomness. Plots for the other measured addresses were similar and are not included here.

---

1 Normally, as noted previously.
To compare the implementation design with our measurements we analyzed the data to determine how many different values each parameter can have. Table 1 shows a count of unique values found in the 11,500 samples taken for each measurement and compares this count to the expected range dictated by the design.

The observed range of stack addresses differs from what is expected but it is clear that our 11,500 samples are not large enough to properly measure the full range and no conclusions can be drawn.

<table>
<thead>
<tr>
<th>Item</th>
<th>Expected</th>
<th>Observed</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stack</td>
<td>$2^{14}$</td>
<td>8,568</td>
<td>-48%</td>
</tr>
<tr>
<td>Malloc$^2$</td>
<td>$2^5$</td>
<td>192</td>
<td>+500%</td>
</tr>
<tr>
<td>HeapAlloc$^3$</td>
<td>$2^5$</td>
<td>95</td>
<td>+200%</td>
</tr>
<tr>
<td>CreateHeap$^4$</td>
<td>$2^5$</td>
<td>209</td>
<td>+550%</td>
</tr>
<tr>
<td>Image</td>
<td>$2^8$</td>
<td>255</td>
<td>-0.4%</td>
</tr>
<tr>
<td>PEB</td>
<td>$2^4$</td>
<td>13</td>
<td>-19%</td>
</tr>
</tbody>
</table>

The occurrences of duplicates in heap and PEB addresses deviate significantly from what is expected. Of 11,500 HeapAlloc address samples, successive samples reported the same value 176 times or 1.5% of the time. If all 32 expected values were equally likely, the probability of successive duplicates would be 3.1% and there would be about 359 duplicates, 51% more than were observed. If all 95 observed values were equally likely the probability of successive duplicates would be 1.1% and there would be only about 121 successive duplicates, 45% less than were observed. This discrepancy suggests that the 95 observed values are probably not uniformly distributed and there may be some bias in the selection of heap addresses when using HeapAlloc. To determine if this discrepancy was due to an obvious pattern, we plotted the number of runs between duplicate values to see if they yielded any obvious pattern, but did not find any (see Appendix IV).

The deviation between expected and observed is even larger for duplicates in the malloc and PEB samples. Duplicates occurred 1.2 times as often as expected in heap

---

$^2$ Heap allocation using the malloc function.

$^3$ Heap allocation using the HeapAlloc function and the default process heap.

$^4$ Heap allocation using HeapAlloc and a heap created with the CreateHeap function.
addresses returned by malloc where 133 duplicates account for 1.2% of the samples rather than the 60 (0.5%) expected if all 192 observed values were equally likely. Similarly, 1,322 duplicates account for 11% of the PEB address samples rather than the mere 719 (6.3%) expected if there were 16 equally likely addresses. These discrepancies point to even larger biases in the randomization of malloc and PEB addresses.

Duplicates occurred 13% less often than expected in the image addresses. This also suggests the presence of bias, although the deviation is smaller than in the other instances.

B. Frequency Distribution Analysis

A distribution that has biases that favor some values over other values is more predictable and has lower entropy. We plotted the distributions of the samples for each parameter we measured to identify any biases.

The distribution of the stack address is shown in Figure 2. The graph shows a near-uniform distribution with no significant biases. Combined with the large range this indicates that the stack addresses should be fairly unpredictable.

![Figure 2 – Distribution of stack addresses.](image)

Significant biases are apparent in the distribution of heap addresses as seen in Figure 3. These parameters also have a much smaller range. As seen in Figure 4, there is also bias in the distribution of code addresses (which are affected by randomization of the image base address) although these biases are not as profound as those seen in the heap distributions.

![Figure 3 – Distribution of heap addresses using malloc, HeapAlloc and HeapAlloc with CreateHeap.](image)

Finally, PEB randomization shows a significant bias to one address. Figure 5 shows that 25% of all executions chose a single address and another 10% of all executions chose a second address. These two addresses were chosen 35% of the time and the remaining eleven addresses were chosen only 65% of the time.

![Figure 4 – Distribution of code addresses.](image)

![Figure 5 – Distribution of PEB addresses.](image)
Because the distributions of image base, heap and PEB addresses have biases, they are more predictable than their range measurements alone would indicate. For example, if all 16 PEB addresses were equally likely, an attacker would have a 6.25% chance of guessing the PEB address but because of biases an attacker can make a single guess that is correct 25% of the time.

IV. CONCLUSIONS

This paper shows that the stack, heap, image and PEB protected by ASLR on Microsoft Windows Vista 32bit RTM have different frequency distributions. While the stack has near uniform distribution over a very wide range, the heap and PEB, and to a lesser degree the image base have much smaller ranges and because of biases in their distributions do not efficiently use this range. As a result, the protection offered by ASLR under Windows Vista may not be as robust as expected. Microsoft has acknowledged the problems with PEB and image base randomization and indicated that it is caused by a weakness in the implementation.

This paper also shows that applications that leverage Microsoft’s HeapAlloc() function are not afforded the same level of protection as those that leverage the ANSI C heap allocation API malloc(). The impact of which is that third party software that explicitly use Microsoft’s API, are potentially more vulnerable from successful exploitation than software that does not. Also apparent is that using CreateHeap() then HeapAlloc() improves the entropy slightly over solely using malloc().

Finally, results show that there are less consecutive duplicates than expected in the PEB randomization. This result adds to the evidence that the source of entropy used within ASLR is poorly used.

ACKNOWLEDGMENT

The author would like to acknowledge the help and support of Oliver Friedrichs, Matt Conover, Zulfikar Ramzan, of Symantec and Tim Newsham of ISEC Partners. The author would also like to acknowledge Nitin Kumar Goel, of Microsoft who reviewed this research and provided candid feedback.

REFERENCES

I. ASLR.CPP

The following code was used to collect the source data for the analysis of ASLR.

```
// (c)2006 Symantec Corporation
// Ollie Whitehouse – ollie_whitehouse@symantec.com
#
#include "stdafx.h"

// Function: Banner
// Description: Print banner
//
void banner(){
    fprintf(stdout,"---------------------------------- ---------------------------------------\n");
    fprintf(stdout,"ASLR - v0.1\n");
    fprintf(stdout,"(c)2006 Symantec Corporation\n");
    fprintf(stdout,"-------------------------------------------------------------------------\n\n");
}

// Prints the location of the function in memory
// This is used to validate .code randomization
//
int verifyCode(FILE *fileCSV)
{
    fprintf(stdout,"%0.8p\n",verifyCode);
    fprintf(fileCSV,"%0.8p\n",verifyCode);
    return 0;
}

// Some of this function taken from Phrack 62
// Prints the location of the PEB in memory
//
int verifyPEB(FILE *fileCSV)
{
    DWORD* dwPebBase = NULL;
    /* Return PEB address for current process
    address is located at FS:0x30 */
    __asm
    {
        push eax
        mov eax, FS:[0x30]
        mov [dwPebBase], eax
        pop eax
    }
    fprintf(stdout,"%0.8X,",(DWORD)dwPebBase);
    fprintf(fileCSV,"%0.8X,",(DWORD)dwPebBase);
    return 0;
}

// Prints the location of a new heap in memory
//
int verifyHeapviaHeapCreate(FILE *fileCSV)
{
    HANDLE hHeap;
    ULONG_PTR *varFoo;
    hHeap=HeapCreate(NULL,1024,2048);
    if(hHeap==NULL){
        fprintf(stdout,"error,");
        return 1;
    } else {
        varFoo=(ULONG_PTR *) HeapAlloc(hHeap,0,100);
    }
```
if (varFoo==NULL){
    fprintf(stderr,"error,\n");
    return 1;
} else {
    fprintf(stderr,"%0.8p,",varFoo);
    fprintf(fileCSV,"%0.8p,",varFoo);
    HeapFree(hHeap,0,varFoo);
    return 0;
}
HeapDestroy(hHeap);
}
return 0;
}

// Prints the location of the heap in memory
// int verifyHeapviaHeapAlloc(FILE *fileCSV)
{
    ULONG_PTR *varFoo;
    varFoo=(ULONG_PTR *) HeapAlloc(GetProcessHeap(),0,100);
    if (varFoo==NULL){
        fprintf(stderr,"error,\n");
        return 1;
    } else {
        fprintf(stderr,"%0.8p,",varFoo);
        fprintf(fileCSV,"%0.8p,",varFoo);
        HeapFree(GetProcessHeap(),0,varFoo);
        return 0;
    }
    return 0;
}

// Prints the location of the heap in memory
// int verifyHeapviaMalloc(FILE *fileCSV)
{
    char *varFoo;
    varFoo=(char *) malloc(100);
    if (varFoo==NULL){
        fprintf(stderr,"error,\n");
        return 1;
    } else {
        fprintf(stderr,"%0.8p,",varFoo);
        fprintf(fileCSV,"%0.8p,",varFoo);
        free(varFoo);
        return 0;
    }
    return 0;
}

// Prints the location of the stack in memory
// int verifyStack(FILE *fileCSV)
{
    int intFoo=1;
    fprintf(stderr,"%0.8p,",&intFoo);
    fprintf(fileCSV,"%0.8p,",&intFoo);
    return 0;
}
int _tmain(int argc, _TCHAR* argv[])
{
    FILE*fileCSV;
    TCHAR strFileName[MAX_PATH];
banner();
    if(argc==2){
_snwprintf_s(strFileName, MAX_PATH - 1, _T("%s\ASLR.csv"), argv[1]);
//fwprintf(stdout, _T("%s\n"), strFileName);
fileCSV=_wfopen(strFileName, _T("a"));
} else {
    fileCSV=fopen("ASLR.csv","a");
}

if(fileCSV==NULL){
    fprintf(stdout,"[!] Couldn't open output file! Exiting!\n");
    return 1;
}

fprintf(stdout,"+-------+--------+--------+--------+--------+--------+
");
fprintf(stdout," Stack | Heap 1 | Heap 2 | Heap 3 | PEB   | Code
");
fprintf(stdout,"+-------+--------+--------+--------+--------+--------+
");
verifyStack(fileCSV);
verifyHeapviaMalloc(fileCSV);
verifyHeapviaHeapAlloc(fileCSV);
verifyHeapviaHeapCreate(fileCSV);
verifyPEB(fileCSV);
verifyCode(fileCSV);
fclose(fileCSV);
return 0;
II. ReSeed.bat

@echo off
REM (c) 2006 Symantec Corp – Ollie Whitehouse – ollie_whitehouse@symantec.com
REM This might seem a little strange but due to some odd instances
REM of not being able to delete/overwrite the executables (although the
REM original process should of finished) because of ‘Access Denied’
REM this is the best way I found to do it without having to put an inefficient
REM 'sleep' in
:Reseed
Copy /Y ASLR.exe ASLR-1.exe
Copy /Y ASLR.exe ASLR-2.exe
Copy /Y ASLR.exe ASLR-3.exe
Copy /Y ASLR.exe ASLR-4.exe
Copy /Y ASLR.exe ASLR-5.exe
Copy /Y ASLR.exe ASLR-6.exe
Copy /Y ASLR.exe ASLR-7.exe
Copy /Y ASLR.exe ASLR-8.exe
Copy /Y ASLR.exe ASLR-9.exe
Copy /Y ASLR.exe ASLR-A.exe
Copy /Y ASLR.exe ASLR-B.exe
Copy /Y ASLR.exe ASLR-C.exe
Copy /Y ASLR.exe ASLR-D.exe
Copy /Y ASLR.exe ASLR-E.exe
Copy /Y ASLR.exe ASLR-F.exe
ASLR=1.exe
ASLR-2.exe
ASLR-3.exe
ASLR-4.exe
ASLR-5.exe
ASLR-6.exe
ASLR-7.exe
ASLR-8.exe
ASLR-9.exe
ASLR-A.exe
ASLR-B.exe
ASLR-C.exe
ASLR-D.exe
ASLR-E.exe
ASLR-F.exe
Del ASLR-1.exe
Del ASLR-2.exe
Del ASLR-3.exe
Del ASLR-4.exe
Del ASLR-5.exe
Del ASLR-6.exe
Del ASLR-7.exe
Del ASLR-8.exe
Del ASLR-9.exe
Del ASLR-A.exe
Del ASLR-B.exe
Del ASLR-C.exe
Del ASLR-D.exe
Del ASLR-E.exe
Del ASLR-F.exe
Goto Reseed
III. Heap Address Selection

The following graph sample shows for heap ASLR the difference in the addresses it selects for each run. The X axis is the run number and the Y axis is the address selected.

We can see from this small sample that there is no apparent pattern in address selection. The above is also true for the complete sample.
IV. RUNS BETWEEN CONSECUTIVE VALUES

The following graph shows the 134 instances where adjacent runs resulted in the same memory address from the `HeapAlloc()` function. With a uniform distribution over a space with $2^8 (=256)$ elements, the probability that two consecutive memory addresses are the same is $1/256$. Based on the results produced by the author we can see the Vista implementation is actually 134 from 11,500, which results in a likelihood of 1 in 85. This result is significantly higher than expected.

The number of runs between these 134 instances were then plotted, the result of which can be seen below. We can see that there was no obvious pattern which reduces the likelihood of an attacker accurately predicting the number of reboots required before the same memory address would be used consecutively.
V. DISTRIBUTION GRAPH - STACK
The following graph shows the frequency distribution of values for each of the stack locations observed during the execution of 11,500 instances.
VI. DISTRIBUTION GRAPH – HEAP (VIA HEAPALLOC FUNCTION)

The following graph shows the frequency distribution of values for each of the heap locations observed during the execution of 11,500 instances. These were used via the HeapAlloc() function, which is a Windows specific API function.
VII. DISTRIBUTION GRAPH – HEAP (VIA MALLOC FUNCTION)
The following graph shows the frequency distribution of values for each of the heap locations observed during the execution of 11,500 instances. These were used via the `malloc()` function, which is a ANSI C API.

![ASLR Heap Memory Location Usage (via malloc)](image)

VIII. DISTRIBUTION GRAPH – HEAP (VIA CREATEHEAP AND HEAPALLOC FUNCTIONS)
The following graph shows the frequency distribution of values for each of the heap locations observed during the execution of 11,500 instances. These were used via the `CreateHeap()` and the `HeapAlloc()` functions, which are Windows specific APIs.

![ASLR Heap Memory Location Usage (via CreateHeap/HeapAlloc)](image)
IX. DISTRIBUTION GRAPHS - IMAGE

The following graph shows the frequency distribution of values for each of the image locations observed during the execution of 11,500 instances.
X. DISTRIBUTION GRAPHS – PROCESS ENVIRONMENT BLOCK

The following graph shows the frequency distribution of values for each of the PEB locations observed during the execution of 11,500 instances.