Holistic Approach for Memory Analysis in Windows System

K. A. Z. Ariffin*, A.K Mahmood, J. Jaafar**, S. Shamsuddin*****
* Department of Digital Forensic, CyberSecurity Malaysia
** Department of Computer Information Science, Universiti Teknologi Petronas
***Department of Research, CyberSecurity Malaysia

Keywords:

Information Forensic
Digital Forensic
Algorithms
Memory Analysis
Signature Search

ABSTRACT

Research on computer memory analysis has been quite intensive in the past years. A number of tools and techniques have been designed and developed to retrieve critical information from the computer memory. However, most of the tools and techniques have their limitation in the ability to retrieve important information. Hence, in the present study, an alternative approach is proposed to combine the process signature search with page table tracking in order to trace all objects that link with the process block. The result from the experiment shows that the new approach is able to retrieve a large number of objects that link with the process block. A good comparison with the previous studies is conducted as to test the efficiency of the new approach.

Corresponding Author:

Khairul Akram Zainol Ariffin
Department of Digital Forensic, CyberSecurity Malaysia,
Level 7, Sapura@Mines,No. 7 Jalan Tasik,
The Mines Resort City,43300 Seri Kembangan, Selangor, Malaysia.
akram@cybersecurity.my

1. INTRODUCTION

With the advent of computer technology, our life has been tremendously affected with both positive and negative impact. The positive aspect of computer technology is in the form of knowledge and information sharing whilst the negative impact is in form of cybercrimes such as online financial frauds, identity theft, impersonating, etc.

The Council of Europe's Cybercrime Treaty defines the cybercrimes as a range of crime that is committed using the computer, network and hardware devices [21]. Further, Symantec divides the cybercrimes into two main perspectives such as the single event that is facilitated by the crime-ware and the range of activity, starting from cyber stalking to stock market manipulation [10]. In RSA 2012 Cybercrime Trends Report, it is reported that the cybercrime has shown no signs of slowing down. The report also concluded that in every minute, 232 computers have been infected with the malware [3].

Due to the increase in cybercrimes, a field known as Digital Forensic has been established. This field involves with collecting, preserving, analyzing, documenting and presenting the contents of computer as evidence of cybercrime [20]. For over a decade, the investigation in Digital Forensics has been focused on the analysis of the non-volatile devices. Until recently, and with the merge of online storage, the volatile device such as computer memory has become critical in the investigation of Digital Forensic.

The research on the volatile memory started in 2005 where Digital Forensic Research Workshop (DFRWS) had organized a Windows Memory Challenge. During the event, two analysis tools had been developed which was known as Memparser [18] and KntList [19]. Nevertheless, password, encryption keys and username were only available in volatile memory which has been listed by Jesse, K [13] as sensitive information where it was used to be stored in temporary processes.A study had been done by Garfinkel, Chows and Rosenblum in 2004 that concluded 86% of the data contents still remained in the computer memory. [22]

An explanation on theories of internal structures and address translation has been outlined in Dhamdhere, M [9] and Russinovich and Solomon [8]. Both explain fully on the function of the internal

structure and address translation in the computer memory with Russinovich and Solomon focusing directly towards Windows Operating System. On the other hand, the procedure of storing data by kernel is also demonstrated by Amari, K [7]. In the demonstration, it illustrates that the kernel has set a pool to store the objects. Most of the data in the computer memory are stored in the paged pool. Carrier, B [23] deduct that most of the data is stored in the paged pool as it allows the data to be transferred into hard disk if the computer memory is running low in space. Meanwhile, the important objects such as process and thread blocks are stored in the non-paged pools as the kernel needs to access them frequently. Since the running process still remains intact all the time if the system still on power, therefore they will be available during the acquisition of the physical memory [15].

Apart from that, Virtual Address Description of a process block has a purpose of tracking the status of the process's address space. This internal structure is maintained by the memory manager and it stores the information on the attributes of the object such as range of the address, inheritance of child and object's security. Due to the information that is stored in this structure, a tool known as VADtool has been developed with the purpose to track the memory mapped files from the memory dump [4][11]. XORSearch [12] is a tool that is designed based on string search technique. It takes a keyword as an input and then performs the search throughout the memory dump. Further, the tool can also help to find the keyword that has obfuscated by using either Exclusive OR (XOR) or Rotate Left (ROL) function that comes with it.

Windows Operating System represents each process in volatile memory as process block. It contains pointers to both next and previous process blocks. Further, the block stores the information on attributes of the process together with pointer to other data structures that is related to it. Process Environment Block (PEB) which is one of the internal structures of process block is responsible to store the location of executable files and the DLL's path. Due to this fact, AccessData [16] group has designed a tool that is known as Forensic Toolkit (FTK). This tool will parse the process block to enumerate all the contents within memory. It also applies Directory Table Base (DTB) information in the address translation algorithms in order to identify the process in memory. Further, Windows Memory Forensic Toolkit (WMFT) has applied this technique by tracking the PsActiveProcessHead link to capture all the active processes in the memory dump [17]. Simultaneously, [5] had demonstrated the data extraction from memory dump by using Kernel Processor Control Region (KPCR). In the demonstration, the information such as running processes, current network connection, file content and other data can be extracted from the image.

S. M. Hejazi, C. Talhi, M. Debbabi [6] has outlined the use of aplication or protocol fingerprint to trace the active application in the memory. The test was conducted to track online application such as email and messenger where from the result, it showed that each of the application had a use fingerprint representation. PTFinder [14] is a tool that applies the file carving where the technique is done linearly to recover only the contiguous file. This technique is applicable because most operating systems will convert the file to be contiguous file instead of fragment files [17].

2. RESEARCH METHOD

The algorithm for the new approach is based on the combination of process signature search and page table tracking. In [1] [2], it is concluded that all the process block in active and inactive mode can be retrieved by using the process signature search technique. Hence, due to this information, this technique will be applied at the beginning of the new approach as the process block is referred to as a critical object and the starting point of the application when it is running in the computer system. Then, once all the process blocks have been retrieved, the algorithm will continue searching for the Page Directory to obtain the Page Table that links with the process block. Finally, the analysis is conducted on the Page Table to capture the address of the important objects that link together with the process block. In summary, the new approach consists of three main algorithms with the purpose of:

a) Track all Process Blocks in Memory Dump

Process block represent a critical object in the computer memory. When an application is opened on the computer system, the kernel will allocate this object in the computer memory. This object will remain in the memory until it has been overwritten by other object or data. Thus, there is a chance that the new exile process block may still exist in the memory. Hence, the mechanism to track the available process blocks has to follow the rules that are discussed below:

RULE 1: Searching for proã with constant hexadecimal value of 50726fe3 (H)

Proã represents the process signature for all process blocks except idle process. It is located outside the process block at a constant offset (in hexadecimal) of 0x01c (for Windows 2000 and XP), 0x0c0 (Windows 2003) and 0x024(Windows Vista). Figure 1 show the location of proã for metasploit.exe in which the application has ended.

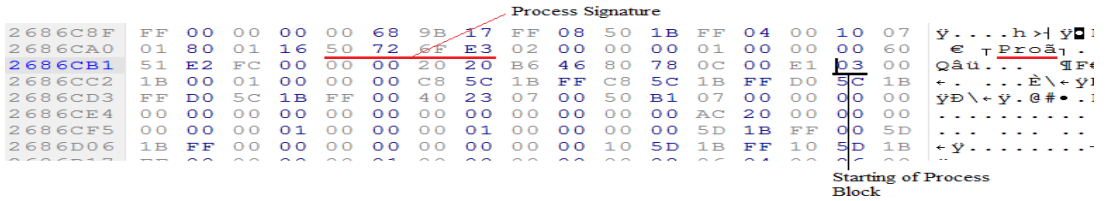

Figure 1. Offset for proã of metasploit.exe block at 0x2686ca4 (Starting of process block at 0x2686cc0)

RULE 2: Searching for true process block

Since proã is located outside the process block, there is a chance that it remains in the computer memory whilst the exile process has been overwritten by other object. Thus, to distinguish the overwritten exile process block with the available one, the value that is stored in ImageFileName is required. ImageFileName is an internal structure in process block that stores the name of the process. In theory, the name for process in the computer system will be in .exe extension. Hence, only the value with .exe extension is chosen as a true entity. Once the true entity has been obtained, it is stored in the database for further investigation.

b) Obtain the Page Table from Page Directory

The next stage of the algorithm is to obtain the Page Table of a process block. For all system architecture, the pointers of the Page Table are stored in Page Directory. Hence, the second algorithm is applied to track the Page Directory of process block (for 32 bit system)

RULE 3: Obtain the DirectoryBaseTable (DBT)

DBT is an important internal structure within the process block. It is referred as CR3 register that plays a critical role in address translation. However, to be precise, in a normal 32 bit system, the value in DTB represents the pointer of the Page Directory. Thus, by obtaining its value it may allow to jump to the starting offset of the page Directory. On the other hand, the 32 bit with Page Address Extension (PAE) and 64 bit systems will require an extra step to obtain the starting point of the Page Directory. It is due to the fact that the value within DBT represents the pointer to the offset of selector for Page Directory. In 32 bit system with PAE, there are four entries for selector in which each of them will store a pointer to individual Page Directory. As in 64 bit system, it is possible to have 512 Page Directories for one process block.

RULE 4: Go to Page Directory and read the pointer to Page Table

Once the Page Directory has been retrieved, the value within it which represents the pointer of Page Table is read. For all system architecture, this pointer is stored in 8 bytes wide format. There are three cases for the value that is stored with the entry of Page Directory (in hexadecimal):

Case 1: all the value is equal to zero (e.g. x00000000)

- This defines that the entry is empty. Thus, the algorithm will move to the next entry.

Case 2: the last number is an even number (e.g. x06785082)

- This defines that the Page Table is no longer in the computer memory. Thus, when this value is found, the algorithm will skip to the next entry.

Case 3: the last number is an odd number (e.g. x07860067)

- This defines that the Page Table exists in computer memory. Hence, the algorithm will store this pointer in the database before moving to the next entry.

The process of reading the pointer of Page Table will stop after the last entry of the Page Directory has been read.

RULE 5: Skip the same pointer of the Page Table

In theory, the process blocks in computer memory will share the same Page Table to store the physical address of linked objects. Therefore, a comparison test is conducted to eliminate the same pointer being stored in database twice.

c) Obtain the address of important objects from Page Table

After retrieving all the data from the Page Directory, the third algorithm is applied to obtain the beginning of the offset for the object in the computer system.

RULE 6: Go to Page Table and read the offset of the object.

All the entries within the Page Table are read to obtain the offset of the object. The mechanism of address storing in Page Table is similar as in RULE 4. It will follow the same case as discussed in RULE 4 where the entry in Page Table could be represented as empty entry, remain and exile object offset. Thus, the same method in RULE 4 is applied to capture the remaining object from the computer memory.

3. EXPERIMENT

In the test, the author uses the holistic approach that is discussed previously to retrieve the amount of object that remains in the computer memory. This test is conducted on a memory dump of 32 bit system that is available on Digital Forensic Research Conference (DFRWS) website [18]

A comparison test between the holistic approach and the previous techniques is conducted. This new approach is tested together with the Process Signature Search (PSS), Process Enumeration technique (PET) and the Hybrid approach to identify the advantages and disadvantages for each of the techniques.

PSS is a technique that uses the object's signature value to capture the processes from the computer memory. However, before the benchmarking test, several experiments have been conducted and the result shows that this technique is also applicable to retrieve other information from the memory image. As an example, the important object that is known as thread can be retrieved by using value of Thrå (Signature for thread). On the other hand, PET is a traditional approach in tracking the object from the memory image where it makes use of Doubly Link between the objects to trace them. The PET is dependent on the address translation algorithm where this algorithm is heavily used for tracing. Finally, the hybrid approach is a combination of both PSS and PET where it has the advantage of tracking the available hidden and exile object from the memory.

4. RESULTS AND ANALYSIS

From the first experiment, the holistic approach is able to trace about 17074 objects that are still in the DFRWS memory dump. In the beginning of the algorithm, 49 blocks have been detected as process block representation. However, only 46 blocks represent a true entity. This is due to the reason that there is duplicate version of the existing process block in DFRWS memory. In this experiment there are three duplicated process blocks as they occur twice in the DFRWS memory. These processes are known as winlogon.exe, dfrws2005.exe and HKserv.exe.Eventually, there are two set of data for the process with a duplicated block. However, the value that is stored in the DirectoryBaseTable (DBT) will remain the same. This result therefore shows that even if the process has two different offset for its two blocks, it has the same Page Directory for both blocks. Due to this information, the objects which are linked with the process block can be traced by capturing their pointers that are stored in the Page Table (pointer of Page Table are stored in Page Directory).

Each of the process blocks has one Page Directory. Hence, there are 46 Page Directories that have been retrieved from DFRWS memory dump. From these Page Directories, about 291 Page Tables have been identified. However, only 12 entities from these pages represent a shared Page Tables as captured from every process block in the DFRWS memory dump. Generally, the shared Page Tables store a huge number of pointers to link object compared to a normal page table. Table 1 shows some of the Page Tables that exist in the DFRWS memory dump.

Table 1.Information on some Page Tables in DFRWS memory dump

Offset of Page Table	Link Process Block	No of active Process
x1031000	All process block	863
x1032000	All process block	553
x103b000	All process block	852
x103c000	All process block	1000
x636000	WinMgmt.exe	4
x83e000	Cmd2k.exe	25
x3ef4000	nc.exe	55
x5fe3000	Asynmgr.exe	18

However, some of the Page Tables in DFRWS memory dump will remain empty. There are two possibilities when the Page Table remains empty:

- Possibility 1

The Page Table acts as reserve to store pointer for new linked objects

- Possibility 2

The empty Page table is for exile process. Since the process has already ended, the kernel removes the pointer within the Page Table as the object is no longer in the computer memory.

Once all the Page Tables have been traced, the location of the object can be captured. Since there is no use of address translation in the holistic approach, the pointer that is stored within the Page Table will not point directly to the object. This pointer will point to the beginning of the page where the object is resided. Hence, there are two methods to obtain the object from this page:

- Do a cross correlation with the result from Process Enumeration technique to allocate the true location of the object
- Scan the page of the object and capture all the recognizable structure

In 32 bit system, the size of a page is equal to 4kb which is equivalence to x1000 in hexadecimal. Thus, when using the second method, the page of the object (normally starts at x000) is scanned from the beginning of the offset until it reaches x1000 (e.g. from x632000 to x633000).

On the second test, the holistic approach is compared with the existing technique. These existing techniques include the Process Signature Search (PSS), Process Enumeration (PE) and Hybrid Approach (combination of PSS and PE). The result from the comparison test is summarized in table 2 while figure 3 shows the number of objects that can be retrieved from all the techniques.

Table 2. Comparison test's result

Factor/Technique	Process Signature Search	Process Enumeration	Hybrid Approach	Holistic Approach
Required knowledge	Operating System	Operating System and System Architecture	Operating System and System Architecture	Operating System and System Architecture
Address Translation	Not apply	Apply	Apply	Not apply
Trace other object than processes	No	Yes	Yes	Yes
Trace hidden and exile processes	Yes	No	Yes	Yes
Trace hidden object besides processes	No	No	Yes	Yes
Number of object retrieved	553*	4101	4695	17075
Number of exile processes	7	0	7	7
Number of hidden object	58**	0	594	12974

*include together with the other objects (e.g threads) that has been retrieved with object's signature
**the linked object with the process block that is still available from the memory image (e.g thread)

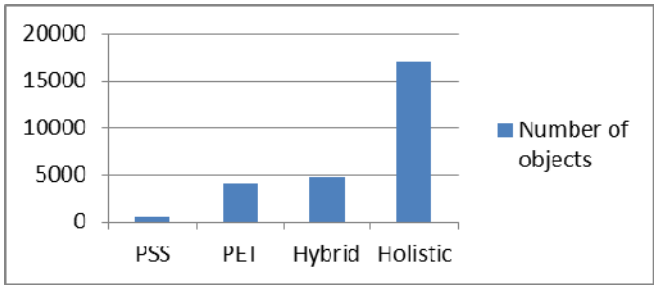

Figure 2. Number of retrievable objects from DFRWS memory dump

Both Process Enumeration and Hybrid Approach require Operating System and System Architecture knowledgesince these techniques will use address translation algorithm. In holistic approach, it does not make use of address translation algorithm. This technique only requires knowledge about Page Directory and Page Table which are covered by system architecture. On the other hand, Process Signature Search does not require system architecture knowledge as it only uses the process signature (proã) to trace all the process blocks in the computer memory. This make it useful due to its ability to trace both hidden and exile process from computer memory as it is involved in both hybrid and holistic approaches.

The Process Enumeration technique has a disadvantage as it cannot detect any hidden and exile object from the computer memory. This scenario happens because the technique cannot trace any object that is no longer in a Doubly Link between active objects. On the other hand, the holistic approach has an advantage against the hybrid since it searches all the entries in the page table. This event does not happen

with the hybrid approach as the process of tracking the object relies directly to the address translation algorithm. Therefore, the hybrid approach is unable to detect some of the entries in Page Table.

5. CONCLUSION

From the result of the experiment, it shows that the holistic approach is able to retrieve more objects from the computer memory compared to the existing techniques. Combining this technique with the Process Enumeration or Address Translation algorithm may result in the data from the object be retrieved directly.For future work, a combination of the holistic approach and the analysis on virtual memory from the hard drive can be experimented to improve on the chance of retrieving other important object.

ACKNOWLEDGEMENTS

The author would like to present his gratitude and appreciation to CyberSecurity Malaysia and Universiti Teknologi Petronas for the help and support that had been accorded during the research period.

REFERENCES

[1] K.A.Z Ariffin, A.K Mahmood, J. Jaafar and S. Shamsuddin, "Hybrid Approach for Memory Analysis in Windows System" WASET 2012 Journal, issue 70, pp 996-1005, 2012.
[2] K.A.Z Ariffin, A.K Mahmood, J. Jaafar and S Shamsuddin, "Process Block Tree (PBT) for Windows Operating System", in ICCEMS 2012 proc, pp.121-128, 2012.
[3] RSA, "The current state of cybercrime and what to expect in 2012", Online Report, 2012.
[4] K.A.Z Ariffin, A.K Mahmood, and J. Jaafar, "Investigating the PROCESS Block for Memory Analysis", in ACS'11 proc, WSEAS Conf, pp. 21-29, 2011.
[5] Ruichao Zhang and Shuhui Zhang. "Windows Memory Analysis Based on KPC", in Proc of the 2009 Fifth International Conference on Information Assurance and Security, IEEE, Xi'An, China, 2009.
[6] S. M. Hejazi, and M. Debbabi "Extraction of forensically sensitive information from windows physical memory", Journal d i g i t a l i n v e s t i g a t i o n vol 6, pp.S 1 2 1 – S 1 3 1, 2009.
[7] Amari, "Techniques and Tools for Recovering and Analyzing Data from Volatile Memory", SANS Institute, 2009.
[8] Russinovich, D.A. Solomon, and A. Ionescu, "Windows®Internals Covering Windows Server® 2008 and Windows Vista®", J. Pierce, Editor., Microsoft Press, 2009.
[9] Dhamdhere, "Operating Systems: A Concept based Approach.", McGrawHill, 1st Edition, 2009.
[10] US-Cert, government organization, "Computer Forensic", Online Report, USA, 2008.
[11] Dolan-Gavitt, "The VAD tree: A process eye view of physical memory", Journal Digital Investigation, pp. s62-s64, 2007.
[12] Stevens, "XORSearch", Unpublished, January, 2007.
[13] Jesee, "Using every part of the buffalo in Windows memory analysis". Journal Digital Investigation, vol **4**, pp. 24-29, 2007.
[14] Schuster, "PTFinder", Unpublished, 2006.
[15] Schuster, "Searching for processes and threads in Microsoft Windows memory dump", Journal Digital Investigation, vol **3,** pp. 10-16, 2006.
[16] AccessData Corporation, " Importance of Memory Search and Analysis", White Paper, Lindon, UT, 2006.
[17] Burdach, "An Introduction to Windows Memory Forensic". Unpublished, Forensic Seccure, 2005.
[18] DFRWS." Memparser Analysis Tool by Chris Betz". Unpublished, 2005.
[19] DFRWS. "Kntlist Analysis Tool by George M. Garner Jr.". Unpublished, 2005.
[20] C. Hill, "What is the Definition of Digital Forensics? ", in eHow, How to do just about everything web page. Unpublished, 2005.
[21] Krone, "High Tech Crime Brief". Australian Institute of Criminology, ISSN 1832-3413, Canberra, Australia, 2005.
[22] Garfinkel, Pfaff, Chow and Rosenblum, "Lifetime is a Systems Problem", In Proc of the ACM SIGOPS European Workshop, ACM, 2004.
[23] Carrier, "A Hardware Based Memory Acquisition Procedure for Digital Investigations", Journal of Digital Investigation, March, 2004.

BIBLIOGRAPHY OF AUTHORS

	Khairul Akram earned his Bachelor and Master degrees with First Class Honours in System Engineering with Computer Engineering from University of Warwick, United Kingdom in 2008 and 2009 respectively. He later joined Universiti Teknologi PETRONAS (UTP) in 2010 to pursue his journey towards academic research and teaching courses to earn his PhD in Information System. During his time in UTP, a number of journal articles and conference papers have been produced and published internationally. Currently, he is appointed as Researcher in Digital Forensic Department, CyberSecurity Malaysia and has been entrusted with the research onembedded system forensics. His passion in research is towards algorithms, embedded system, image processing and audio authentication. He is a member of IET professional group.
	Ahmad Kamil Mahmood earned his Bachelor and Master degrees in Actuarial Science and Statistics from the University of Iowa, Iowa City, USA in 1986 and 1988 respectively. He later joined UUM,Bank Negara Malaysia, and Public Service Department andPETRONAS. After 10 years in the industry, he continued his journey in the academia serving the Universiti Teknologi PETRONAS in 1998 teaching courses for the Bachelor Degree inICT and BIS. He earned his PhD in Information Systems from the University of Salford, UK in 2005. With his research team, a numberof journal articles and conference papers have been produced andpublished internationally. Currently, his industrial collaboration research projects keep him occupied while supervising 7 postgraduate students and assuming the Dean of Faculty of Science andIT.
	Jafrezal Jaafar obtained B.Sc in Computer Science from Universiti Teknologi Malaysia in 1998, MAppSc. (IT) from RMIT University (Australia) in 2002 and PhD from University of Edinburgh (Scotland, UK) in 2009. Previously he works as System Engineer for several years. He is currently the Head of Department for the Computer & Information Sciences Department, Universiti Teknologi PETRONAS, Malaysia. His research interests are in the area of Soft Computing and HCI. He is actively involved in a number of research works and secured research grants in these areas. He has also produced numerous journal, conference and workshop papers.
	Solahuddin received his PhD from University of Bradford, United Kingdom in Network Security in 2008. He received a post-graduate Diploma in Systems Analysis from UiTM in 1991. He started his career with the Malaysian Armed Forces after completing his first degree in Electrical Engineering from Wichita State University, USA in 1986. He served in the Royal Signal Regiment of the Malaysian Army for 10 years holding various posts as communications engineer and IT manager before joining the industry after the completion of his stint with the Malaysian Armed Forces. In 1997 he joined Softlabs Technologies Sdn Bhd as the General Manager. He was entrusted to manage and lead system development projects with various industries such as oil and gas, defence, telecommunications and local governments. In 2002 he joined National ICT Security & Emergency Response Centre (NISER) now known as CyberSecurity Malaysia as the Expert Service Manager. Later on, he was entrusted to be the manager forMalaysia Emergency Response Team (MyCERT). He has earned 4 professional certifications namely CWNA, CISSP, CEH and BS7799 Lead Auditor. With his knowledge and skills in various security domains, he is now entrusted to be the Chief Technology Officer at CyberSecurity Malaysia. He is also the research coordinator for CyberSecurity Malaysia.