Research on New PE File Packer and Shelling Methods

Xiaoyu Liu [1, a], Haichen Zhu [2,b]

[1] Chongqing University School of software, Chongqing400060, China

[2] Chongqing University School of software, Chongqing400060, China

[a]pv3@qq.com, [b]121495580@qq.com

Keywords: PE file, Packer, Shelling, Virtual machine

Abstract. The full name of PE file is Portable Executable file. The common EXE, DLL, OCX, SYS and COM documents are all PE files. File packer is a necessary means of application authors usually use to protect copyrights, but it can be used by many malicious softwares to avoid the detection of anti-virus softwares. Common shelling softwares usually deal with these programs by finding the feature codes of the targeted packer files, while directional shelling softwares usually find by specified features which have already been concluded by Network Security engineers, However, with the development of shell protection, more and more shell applications can't be processed by common shelling softwares as well as directional shelling softwares .To solve the threat of these malicious softwares, new shelling methods must be developed. The paper introduces new shelling and packing ways, and focuses on introducing principals and applications of these techniques.

Introduction

The PE file is an executable file format of windows, which includes EXE and DLL. Encrypting PE files can prevent source programs from being cracked or decompiled, so packer technique is very important for protecting the copyright of authors. However, with the development of packer technique, more and more Trojans use it to make it not be detected. For Network Security Engineers, the core and soul to deal with malicious programs is analyzing and then attempting to decryption as well as shelling, which is very important as a standardized process. So in this way, the development of shelling technique has great significance for Network Security Engineers. Based on the conventional shelling techniques, the paper analyzes and makes the research of new shelling ways, and proposes the prospects and applications of these algorithms. Lastly, the paper introduces counter ways for new shelling algorithms, and introduces principles and applications for the newly invented algorithms.

Introduction of PE File Structures and Basic Knowledge of the Shell

The paper introduces the basic structure of PE file, and introduces the concept and classification of shell.

PE file structure. The structure of PE file is as follows shown in Fig. 1.

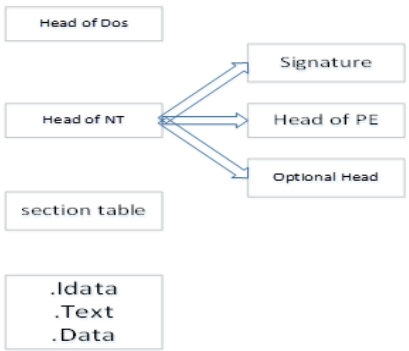

Fig. 1 Analysis of PE file structure

The important parts of the figure are introduced below: DOS stub is compatible with MS-DOS operating system, the objective of which is to prompt a text when the file operates on MS-DOS, such as the recommendation to customers: This program cannot be run in DOS mode, and it also indicates the position of files we will next analyze. When users see these recommendations, they will know that these files cannot be operated in windows system,

NT header includes important information of windows PE files, which includes a signature information with PE, and MAGE_FILE_HEADER, in addition with IMAGE_OPTIONAL_HEADER32.

The Basic Concept and principal of The Shell. The shell usually contains specific codes. In packer applications, the shell is operated before real application, and it can firstly get the control immediately. After the shell detects that the external environment, and the operation environment is confirmed to be safe, the real code can be executed after being decoded. Packer uses the special codes to compress the executable files or convert the data. Packer softwares are generally divided into two types:

1. Compression shell. The objective of the packer software is not to encrypt but to save space greatly. The principal of the packer software is similar to the common file compression program. After compression, the size of PE file is reduced greatly, and GZIP and LZ77 are preferred when choosing packer algorithms.

2. Encryption shell, and it is the object of study in the paper. The packer software doesn't consider the size of the occupied space, but it concerns the safety and complexity of encryption. By confusing the data structure and key part of original PE file, the shell makes it more difficult to crack and decompilation.

When the packer program executes itself, the packer part would execute at the first time when the program is loaded into memory. When loaded , the packer part will decode the code of the program into already decrypted one and map these codes to memory.

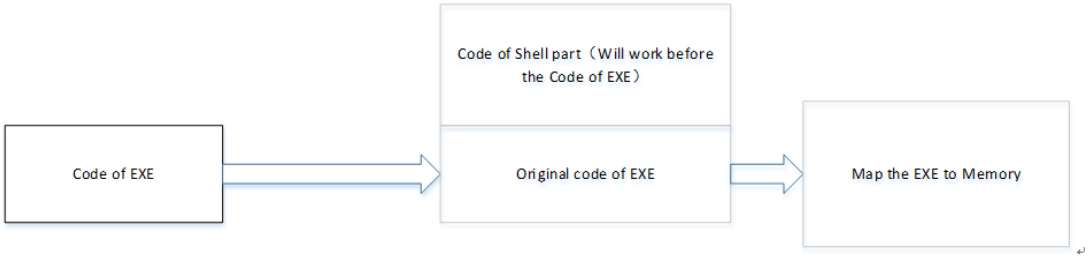

Fig. 2 Execution of packer application

Artificial Shelling Principals and Conventional Shelling Algorithms

Artificial shelling principal. The common shelling process is as follows:

Scanning shell (PEID、FI、PE-SCAN)--->finding OEP(OD)--->shelling/Dump(Lord PE、PeDumper, shelling plug-in unit of OD and PETools)--->repair (Import RE Constructor)

Generally speaking, the scanning shell operation means using codes to judge if the application is packed. Finding the OEP is to find shell codes and skip code segments. After decompression and reduction, the shell program is executed on OEP, and the memory mapping file is the decompressed program. At this time, the memory mapping file can be captured. (The process is called Dump). And it is not necessarily captured at the original entry point of the program, as long as the memory mapping file is ensured to be restored. Repairing the IAT is the last step in the shelling process, and the essence is to decode and repair the broken IAT.

Finding OEP is the most important, and is the core and soul of artificial shelling methods. From the analysis, we can get that the codes of programs described by different languages at OEP are as follows.

Table 1 OEP feature points in different languages

Language	Machine Code	Assembly Code
Delphi	55 88EC 83C4 F0 B8 A86F4B00	PUSH EBP MOV EBP, ESP ADD ESP, 10 MOV EAX, PE. 004B6FA8

VC ++	55
	8BEC
	83EC 44
	56
		PUSH EBP
		MOV EBP, ESP
		SUB ESP, 44
		PUSH ESI

VB	FF25 6C104000 68 147C4000	JMP DWORD PTR PUSH PE. 00407C14 CALL < JMP. &MSVBVM60. # 100 >
DASM	6A00 E8C50A0000	PUSH 0 CALL<JMP.&KERNEL 32GeModu H andA>

Analysis of directional shelling tools and common shelling tools. Directional shelling tools are developed for a certain kind of shells. The characteristic value of the objective packer tool has been known, so the speed to decode is faster than the speed of common shelling tools. Some antivirus softwares use directional shelling tools, the example is ClaimAv.The directional shelling tool is the best if judged only by speed, but it has evident disadvantages. Firstly, for each type of shell, a corresponding shelling algorithm should be singly written. Secondly, the directional shelling tool must have lots of feature-based algorithms, which make shelling softwares occupy more spaces. In recent years, various new packing ways have appeared, if using directional shelling tools only, it probably cannot meet the requirement for decoding ,Especially for anti-virus software, if too large space is occupied , it must influence the operation efficiency and user experience.

The common shelling tools include RLdePacker 1.3. Compared with the characteristics of directional shelling tools which are lack in universality and difficult to be classified as well as needing artificial analysis, common shelling tools not only can reduce manpower cost, but also have higher shelling efficiency for unknown virus compared with directional shelling tools.

However nowadays, as the development of packing technology, algorithms now available cannot meet the requirement for decryption, so it is urgent to develop new shelling algorithms.

Introduction and Research of New Common Shelling Algorithms

Automatic shelling technique based on virtual machine. The innovation of the technique is that all modules are operated on a virtual machine. Under the condition that the virtual machine is safe and stable, and can execute the orders and requirements rightly, the shelling application is hard to be found by the packer application.

The standardized module is to restore the confused program codes by mixing with lots of garbage codes. The function of the module is to extract and analyze the decoded codes of object programs when object program dynamically decodes the encrypted parts during their runtime

The program reconstruction module includes position and repair of IAT and reconstruction of input list. After processing reconstruction module, the program becomes a completed executable file.

The method has evident disadvantages, which are listed as follows.

1. The premise that the method is reliable is that it is established on the basis of stable virtual machine. If the virtual machine is not reliable itself, or the operation time of doing these instructions and the efficiency for these instructions have significance difference from the mainframe, then it will be more likely to be found by the reverse program. If detected, the method would be ineffective.

2. Packer applications must once showed integrated decoding codes in the memory during the runtime cycle of the program. With the development of encryption shell protection, more and more encryption tools and programs use new confused algorithms, which makes shelling more difficult.

Shelling algorithm based on environmental sensitivity analysis. The principal of the shelling method is shown in Fig. 3

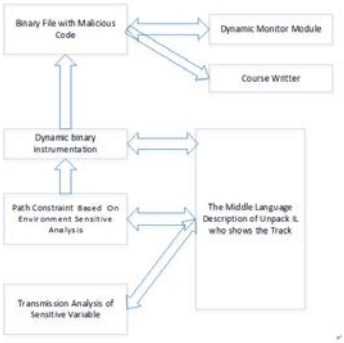

Fig. 3 Automatic common shelling algorithm based on environmental sensitivity analysis

Unpack IL is a static single assignment intermediary language including assignment statement, assertion statement, jump statement and condition jump statement.

Malicious codes need to leak the path information relating to environment when it is executed. The method uses the effective inference to clear the environmental sensitivity and guide the malicious codes to execute the shelling process. Compared with other shelling techniques, this method can effectively resist the interference of environmental sensitivity technique. And the shelling efficiency is better than the common shelling software.

However, the method has evident defects which are remained to be solved. On one hand, if the environmental relationship of malicious code is simple, using shelling applications based on sensitivity has no evident advantages compared with common shelling softwares. On the other hand, if packer application is inclined to enhance path protection, for example, path confusion or other technical means are used to prevent from path being leaked, then in this condition, the effect of the technique is not evident.

Analysis and Research of New Object Code Confusion Algorithm

To fight against these shelling algorithms, several new algorithms are proposed to face the threat of shelling algorithms.This paper analysed and Studied the New Object Code Confusion Algorithm for example.

The algorithm uses irreversible-control confused ways to modify the transfer instruction of program source code. When the transfer instruction operates, it must go through shells at the beginning, which makes it more difficult to track transfer instructions. After reading assembly results, the new destination address will be computed, and the source program will be modified, in the end, new confusion data finally generates.

According to the theory of Collberg C, the algorithm can be divided into layout confusion, data confusion, control confusion and preventive confusion.

Layout confusion. It disturbs the layout of program and uses specific characters to replace source codes. For example, after executing layout confusion, the source program is replaced by the new disordered program. For example, if the original program is just as follows:

String A=“TEST”

After executing confusion algorithm, the program may become 0ooo000TEST, and it is more difficult to crack.

Data confusion. Data confusion modifies the data filled with PE file rather than code filed.

Control confusion. Control confusion uses fuzzy predicate and embedded outreach to make crack more difficult by breaking the space distribution of codes.

Preventive confusion. The application of preventive confusion needs to combine the disadvantages of specific decompiler. For example, the decompiler Mocha doesn't translate the instructions after instruction Return. Directed against the defects of some decompilers, specific ways of encryption can be used for prevention. For example, Mocha can be used to put instructions after return to avoid decompilation.

Limited by the length of the article, the paper only introduces one method. Elliptic curve shelling way, system infusion encryption and the way of dynamic technique encryption are also used in actual application.

Conclusion and Prospect

The paper firstly introduces PE file format and principle of shell to make the reader know preliminary understanding of shelling and packing. Then, the paper introduces the basic knowledge and technological process of artificial shelling, and introduces two common shelling methods: the directional shelling method and the common shelling method. The writer focuses on describing the characteristics and defects of two methods, and introduces the technical principles of the methods.

After studying the existing shelling techniques, the paper introduces two new algorithms, and analyzes the decoding algorithms, and proposes the principle and features.

In order to fight against these new shelling techniques, the writer lastly introduces a technique which improves the existing encryption technique, and proposes that the obvious advantage compared with the common shell-encryption technique.

With the development of packer technology, the future shelling technique will be more associated with operation system, and encryption algorithm will be more complicated. Shelling technique still needs to be improved and innovated.

References

[1] Z. Wang, C.F. Jia, K. Lu: Chinese Journal of Compters., Vol.35(2012) No.4, P.693.

[2] R.F.Tian, X.F.Wei: Market Mordernization, Vol.532(2008) P.1.

[3] X.X. Peng, Z.J. Hu, T. Gong, H. Shu:Technology Research, Vol.5(2014) P.1.

[4] L. Li, Q.J.Liu and D.R.Xu: Computer Applications and Software, Vol.27(2010) No.9, P.279.

[5] D.l. Yang:Technology Apply, Vol.4(2013)， P.55.

[6] L. Li: Design and Analysis the Key Technologies of Packing Using by Object Code Obfuscation (MS., Su Zhou University, China 2009) p.1.

[7] S.C. Yu:A Universal Automatic Unpacking System Based on Virtual Machine(MS.,University of Electronic Science and Technology of China,China 2010) p.1.

[8] Z.S.Zhao: The Design and Implementation of Unpacking Platform in Windows(MS., University of Electronic Science and Technology of China, China 2012) p.1.

[9] Information on http://www.onegreen.net/Article/HTML/598.html(2014.12.1).