

# The Protection of Intellectual Property of Computer Software Based on Computer Forensics

Wu W<sup>1</sup>, Mai Y<sup>2\*</sup>, Zhang J<sup>2</sup> and Zhou J<sup>3</sup>

<sup>1</sup>Department of Computer Science, The University of Hong Kong, China

<sup>2</sup>Department of Information Technology, Hubei University of Police, China

<sup>3</sup>Department of International Policing, Hubei University of Police, China

**\*Corresponding author:** Yonghao Mai, Department of Information Technology, Hubei University of Police, Qiaokou District, No.99 Nanniwan Avenue, Wuhan, Hubei Province, China, Tel:+8618171519603; Email: myh9999@163.com

## Perspective

Volume 4 Issue 1

**Received Date:** January 08, 2019

**Published Date:** January 22, 2019

**DOI:** 10.23880/ijfsc-16000156

## Abstract

As the proportion of computer forensic in the cases of intellectual property protection of computer software is increasing, the extraction and analysis of electronic evidence will directly affect the result of the case. However, from the acceptance of the case to the final opinion book, there are widespread problems in forensic processes, techniques, and the laws and regulations in cases involving intellectual property protection of computer software, which lead to the slowdown even termination of the case. The paper studies on laws and regulations, forensic process and technique of computer, and comes up with solutions to these problems.

**Keywords:** Computer Forensic; Computer Software; Intellectual Property Protection; Judicial Identification; Electronic Evidence

## Introduction

The popularization of the Internet has brought tremendous impact on people's lives. According to relevant data, the number of Chinese Internet users has reached 751 million by June 2017, the penetration of Internet is 54.3%, and the proportion of mobile Internet users accounts for 96.3% [1]. The diversification of the Internet's transmission methods and its low cost make it possible for infringers to obtain profits quickly. Therefore, how to fix evidence through computer forensics has become a very critical issue.

According to data released on the official website of the Supreme People's Court of the People's Republic of China, there were more than 50,000 first-instance cases of

intellectual property infringement in 2015, and the number reached more than 70,000 in 2016, an increase of 41.34% compared with 2015. There were as many as 28 countries involved in intellectual property infringement cases, of which the United States, France, and Germany involved the most. In terms of the trial period, the infringement of computer software copyright exceeds 150 days, 105 days more than the average [2]. It is obvious that there already exists the problem of long trial period and difficulty in safeguarding property rights. This article clarifies its difficulties from the three aspects of the relevant legal provisions, forensic processes and the forensic techniques, and puts forward suggestions for improvement.

On January 2nd, 2018, the Supreme People's Court issued the "Notice on Giving Full Play to the Judicial Function to Create a Good Legal Environment for Entrepreneurs in Innovation and Entrepreneurship" [3]. It emphasized the need to protect the intellectual property of entrepreneurs in accordance with the law, and to solve the problem of low costs in infringement on rights but high cost in protecting them. On February 28th, 2018, the General Office of the CPC Central Committee and the General Office of the State Council issued the "Opinions on Several Issues Concerning the Enhancement of Reform and Innovation in the Field of Intellectual Property Judgment", focusing on solving problems such as "difficulty of proof, low compensation, and long period" [4]. It can be seen that the country is accelerating the construction of the legal system of intellectual property and the gradual improvement of the trial system for intellectual property rights.

After four years of preparation and debate the GDPR (General Data Protection Regulation) was finally approved by the EU Parliament on 14 April 2016. It was enforced on 25 May 2018 – and organisations that are not compliant could now face a heavy fine, which is aimed at protecting all EU citizens from privacy and data breaches in today's data-driven world. Although the law about personal information protection have not been published yet, Chinese government is working hard to enforce it.

### The Laws on Cases of Infringing Computer Software Intellectual Property

Intellectual property is the intellectual achievement of human-beings' creative labor. It is actually an expression

of social credibility that inspires people innovate boldly and trustingly, and even more it is a representation of resistance towards hypocrisy, ugliness and evil [5]. Intellectual property include patents right, trade mark right, and copyrights. The "Criminal Law of the People's Republic of China" stipulates the crime of infringement of intellectual property related to computer software as the crime of copyright infringement in "infringement of intellectual property" [5] (Figure 1).

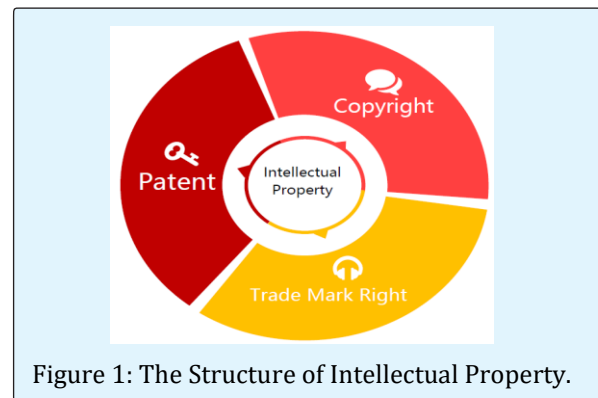


Figure 1: The Structure of Intellectual Property.

The "Opinions on Several Issues Concerning the Application of Law in Handling Criminal Cases of Infringement of Intellectual Property" and the "Provisions of the Ministry of Public Security of the Supreme People's Procuratorate on the Economic Crimes which are handled by Public Security Organs" clearly stated that when a case-handling organ handles a criminal case of infringement of intellectual property, if there are any items that need to be identified, they should entrusted with the expert testimony institutions that have been certified by the state [6,7] (Figure 2).

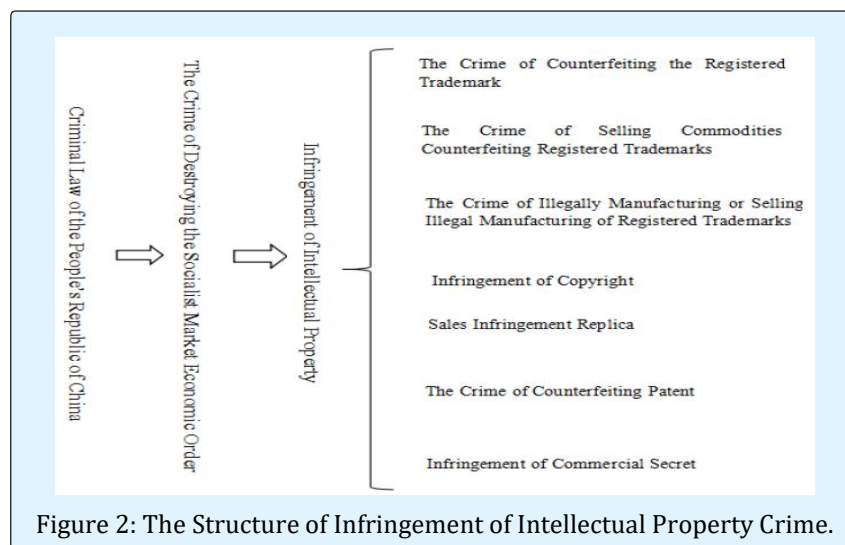


Figure 2: The Structure of Infringement of Intellectual Property Crime.

According to the “General Principles of Expert Testimony Procedures” promulgated by the Ministry of Justice, the expert testimony procedures are divided into four major steps: the entrustment and acceptance of expert testimony, the enforcement of expert testimony, the issuance of opinion books, and the appearance of judicial appraisers in court. Article 11 of the second chapter stipulates that the expert testimony institutions shall uniformly accept the expert testimony entrustment of the case-handling organ which refers to the investigating organ, the examination and prosecution organ, and the judicial organ that handles the lawsuit [8].

### Computer Software’s Similarity Identification

#### The Relationship Among “Similarity”, “Identity”, “Substantial Similarity”, and “Same Expression”

In the practice of computer software similarity identification, the four expressions “similarity”, “identity”, “substantial similarity” and “same expression” appear frequently and it’s so confusing. The author believes that “similarity” is somewhat a hyponymy of the other three as the similarity of software means that both “form” and “idea” are similar while the meaning of “identity” and “substantial similarity” is the same, both indicating that the “kernel” of the software are identical. The most convincing thing is the comparison of the software source code, if the idea, structure, and expression of the sample source code are similar or the same, it can be said that they have the same identity, excluding third-party developed code, third-party library files, and the public part. “Same expression” means the same “form” of the

software, including the software’s directory structure, interface, operating functions, operating results, installation procedures and so on. Therefore, only when the “form” and “idea” of the software are similar or same, can the two software be called “similarity” (Figure 3).

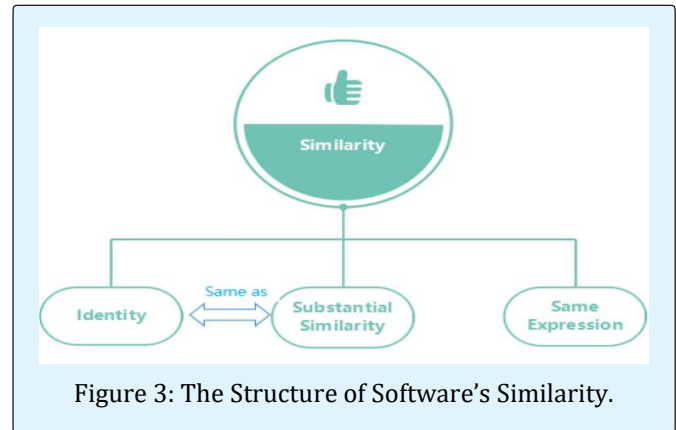


Figure 3: The Structure of Software’s Similarity.

#### The Process of Identification

According to the official document issued by the Ministry of Justice, computer similarity identification already has a standardized operating procedure, which includes seven steps: records of samples and test material, backup of samples and test material, selection of inspection items, comparison of procedures, comparison of documents, inspection of records, and inspection of results (Figure 4).

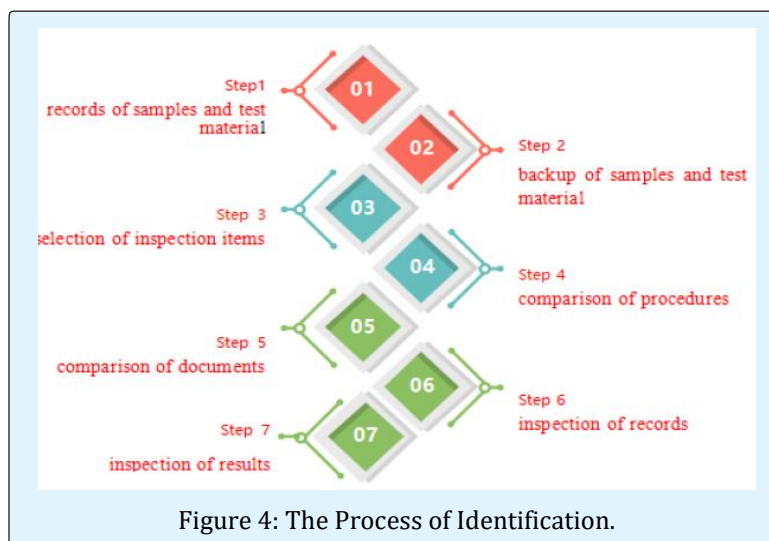


Figure 4: The Process of Identification.

Among them, the third step that is the selection of inspection item, is indicated in the official document that the source code, the object program, and the document need to be compared with each other [9]. The comparison of the source code emphasizes that of the “kernels” such as the ideological logic and frameworks of the core program, while the comparison of the object programs and documents emphasizes that of the intuitive “facade”. In practice, the source code of the software is often the most difficult part to obtain, especially the source code of the sample program when reverse engineering technology is needed. However, in order to protect the intellectual property in the production of software and to prevent reverse analysis, entrepreneurs provide “packaging” protection for the software whose source code cannot be analyzed by the general reverse software thus disturbing the acquisition of electronic data.

Secondly, the author believes that the object code can also be added into the inspection item. Object code refers to the code generated by a compiler or assembler in computer science after processing source code. It is usually composed of machine code or code that is close to the machine language [10]. Object code is a kind of low-level language that can be recognized by computer. It is formed by high-level programming language through compiler conversion. Therefore, strictly speaking, it is the compiler instead of the person that directly forms the object code. So it is the “form” of software rather than the “idea” that is reflected by the comparison of object code.

### The Relationship Among “Source Code”, “Object Code”, “Object Program”, and “Document”

“Source code” refers to the uncompiled text file written in accordance with a certain programming language specification. And it is a series of human-readable computer language instructions [11]. The characteristics of “source code” are human-readable. The appraisers can understand the thoughts and logic reflected by the software through reading the source code, and then determine whether the software has the identity; “object code” refers to the computer-readable machine instructions after the source code has been compiled by the compiler. “Object code” is characterized by computer-readable, it cannot be directly understood by human subjectivity, and the appraiser needs to use professional software to compare with “object code”; “Object program” means the source code is compiled object code set that can be directly run by the computer [12]. The “object program” is characterized by being run by the computer. The appraiser can view the software’s function, record the

software installation process, the software operation interface, and other objectively visible content by running the “object program”; the scope of “document” refers to not only textual content during the writing of the software, but also developing documents, requirements specifications, design plans, marketing plans, patent books, and other documents which from the development of the software to be reserved as a test material.

Based on the above introduction, it is not difficult to understand that “source code” has been compiled to produce “object code”, the set of “object code” forms the “object program”, and “object program” is one part of “document” (Figure 5).

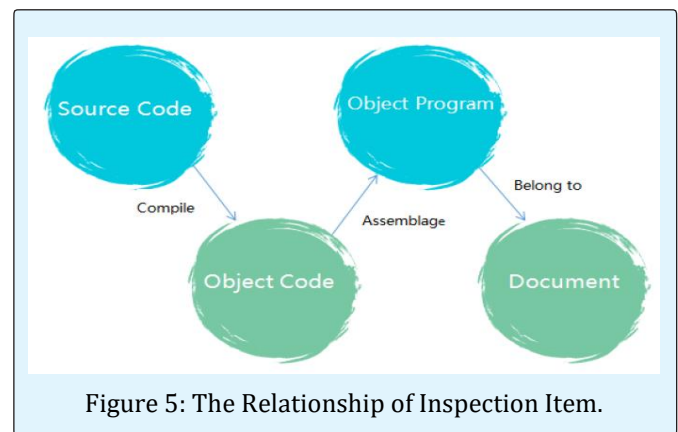


Figure 5: The Relationship of Inspection Item.

## Comparison Technique of Computer Software Similarity Identification

### Comparison Technique of Source Code

In the software similarity identification, the comparison of source code is indispensable. Source code is designed independently by the programme, and it is also the electronic data that represent the thought value of software most. The analysis of source code can help understand the idea of software design from the core ideological level. Each software is designed to satisfy the needs of the public as where there are needs, there is market. We can quickly capture the features of the software and the differences between the software and others by analyzing the part which is designed to achieve the goal. The differences are the most valuable and venerable part of the software, and it’s also the most important part to identify software similarity.

However, general software has as large as thousands of lines of source code How to quickly identify the core part from the thousands of lines of code has the following

requirements, firstly information from the software designers, secondly professional comparison tools, and thirdly certain comparison skills.

For the professional tools, "Beyond Compare" is recommended, which is a set of file comparison tools

introduced by Scooter Software. The main purpose is to compare two folders or files, including the local file and the remote file. The red part indicates the difference, and the black indicates the similarity (Figure 6).

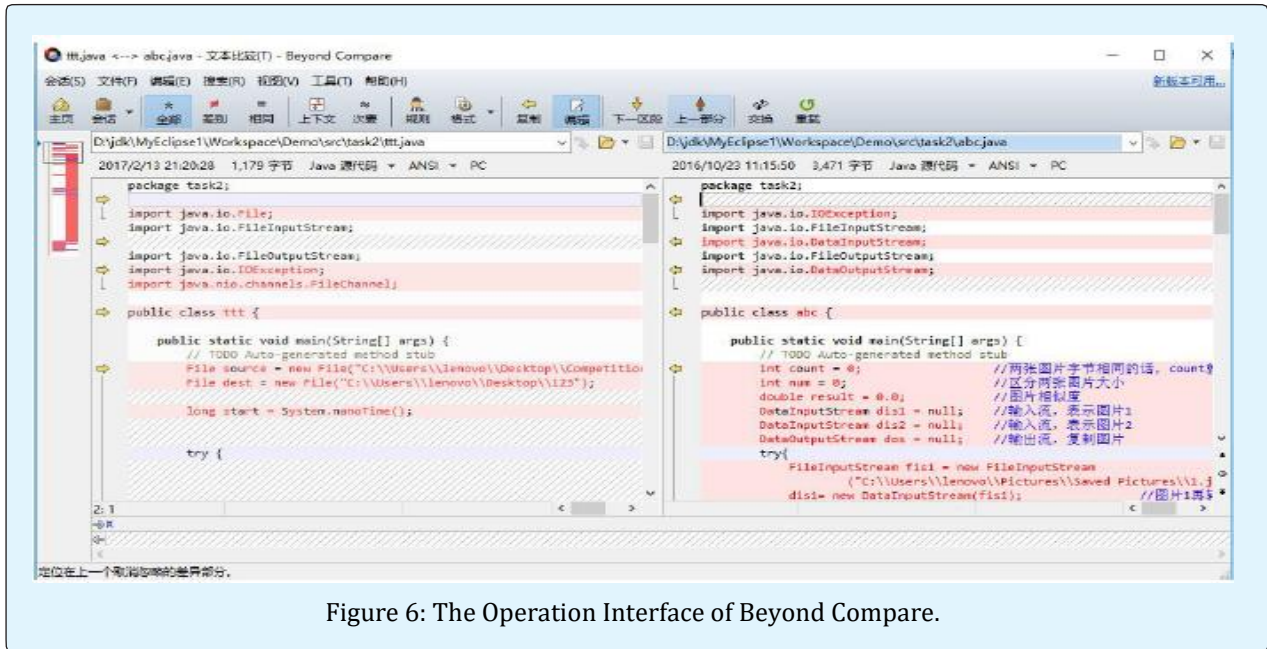


Figure 6: The Operation Interface of Beyond Compare.

The instructions of FC and COMP in DOS can also implement the function of comparing two files (Figure 7 & 8).

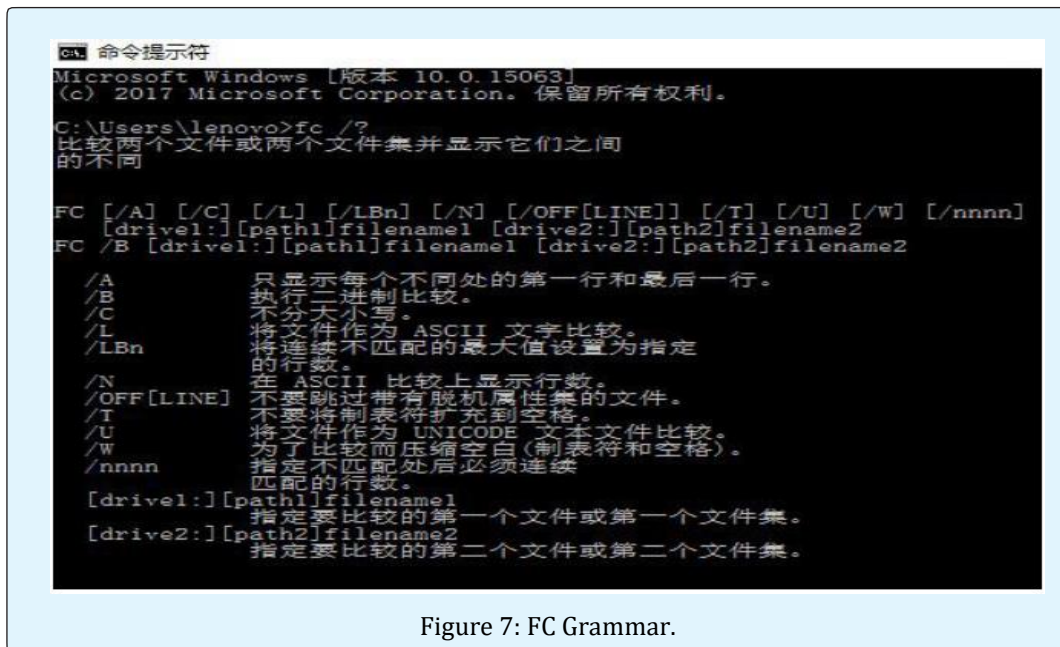


Figure 7: FC Grammar.

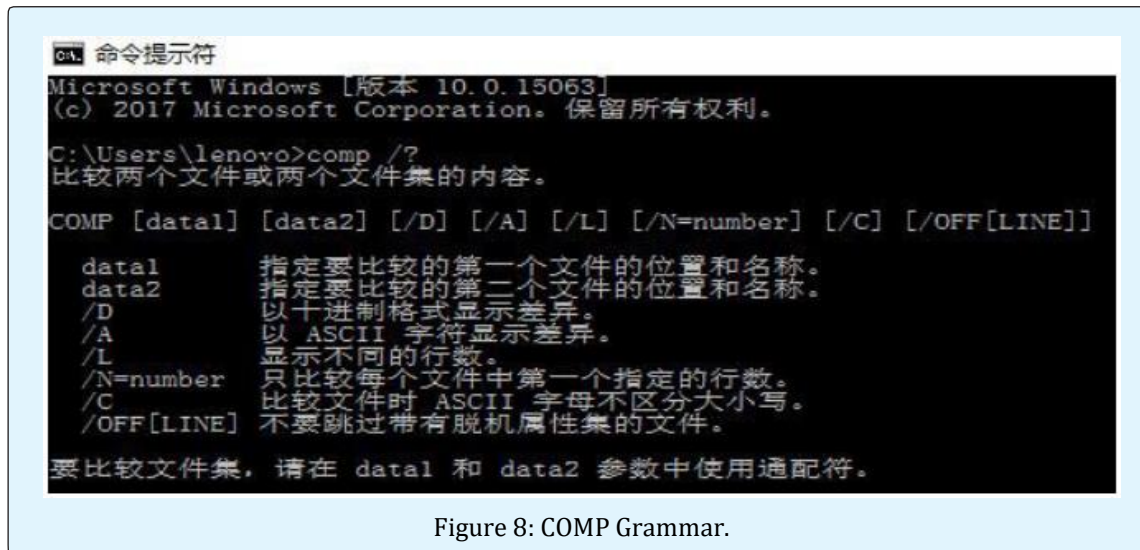


Figure 8: COMP Grammar.

As we all know, program writing is subjective. A software written by different programmers, though the function is the same, differs in the source code. Because each programmer has its own programming habits, it is personalized in terms of certain grammar calls, program layout, comment content, and code modifications. Comparison of individual parts tells whether the source code of the software is the same at the “idea” level.

### Application of Software Reverse Engineering

In the process of practical forensic investigation, because of some objective and subjective reasons, the

source code of the sample is not always available. The source code of the test materials is provided by the infringer that is when the reverse engineering technology is needed to acquire the source code of the sample. Reverse engineering is a technology to obtain software code compared to software development. Traditional software development is the object program that the source code is generated by the compiler and can be executed by the computer. The reverse technology is a technology that resolve the corresponding source code by analyzing the object program [13] (Figure 9).

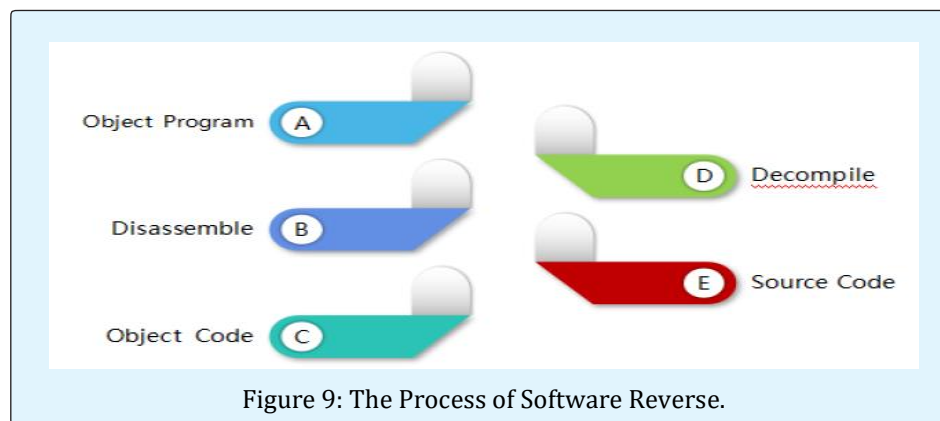


Figure 9: The Process of Software Reverse.

Software reverse can be divided into direct reverse and debugging reverses [14]. Direct reverse means that the reverse tool is used to disassemble and decompile the

object program to obtain the source code without running the program. After the reverse succeed, it analyze design idea, algorithms, grammar, and other aspects of the

program; and debugging reverse indicates that the program is debugged during the reverse process, so that it is convenient to observe changes in the values of function variables and registers when the program is running. In practice, it is often the combination of two methods that can find the object more efficiently.

Not all the software can restore its source code in a reverse method. Some software is designed to prevent software code from being reversed by outsiders, so they

add protective actions by using hardware or software. This is so-called technology of “packaging”. This type of software that has been added protective actions is very difficult to reverse. However, the action of encrypting and protecting software is more likely to happen to the infringed party, and the infringer prefers to handle the software in a low-cost way in order to obtain profits, while the cost of the “packaging” technology is too high and does not meet the interests pursuit of the infringer (Figure 10).

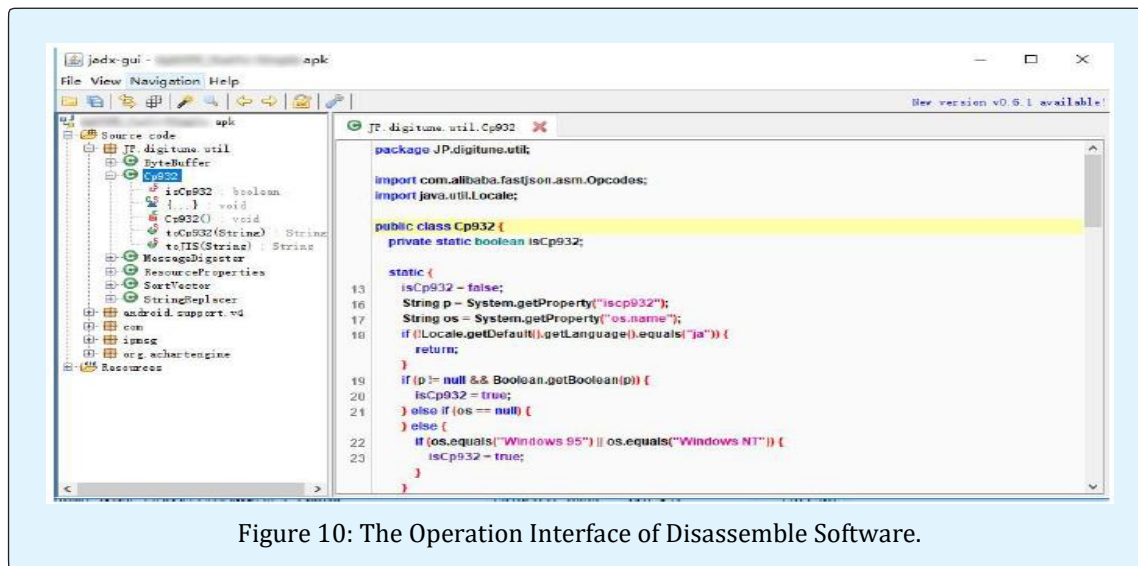


Figure 10: The Operation Interface of Disassemble Software.

## Hash Check

In the expert testimony, the hash value has the effect of judging whether the test material before and after the identification was changed. In the identification of the software intellectual property, it can be used as the basis for determining whether the sample and the test material are consistent. In the official document released by our country, it is pointed out that if the hash value of the object program file corresponding to the sample and the test material is the same, the software has the identity [15]. The hash function is characterized by unidirectionality and collision constraints. Unidirectionality refers to the irreversibility of its operation direction. In the hash function, it means that only the output can be derived from the input, and the input cannot be calculated from the output. The collision constraint means that an input cannot be found that make its output is equal to a known output result, or two different inputs cannot be found at the same time so that the output results are exactly the same. A function only

satisfies these all characteristics strictly at the same time, can we recognize such a hash.

In the handling process of practical case, MD5 value and SHA1 value are commonly used, but these two algorithms have been proved to be insecure. As early as in 2004, Chinese mathematician, Wang Xiaoyun, proposed an algorithm that can successfully break through MD5. The team of Google and CWI released the first SHA-1 collision on February 23, 2017. At present, the algorithm above SHA256 should be used which is considered safe. The author believes that MD5 and SHA1 should be deleted and no longer used in the “Implementation Regulation of Software Similarity Identification”.

## Conclusion

This paper starts from the perspective of computer forensics, and studies some issues that arise in the protection of intellectual property in computer software. At the legal level, the contradiction in the judicial process

between the case-handling organs and the expert testimony institutions is a problem demanding prompt solution; at the forensic level, several confusing concepts have been distinguished; at the technical level, the three key technologies involved have been explained, and the technical loopholes have been pointed out. Based on the above study, there must be some omissions. The author hopes that the future research about this issue could be more comprehensive and thorough.

### Acknowledgements

This study is supported by the Major Project of the Ministry of Education of China (Grant No.17JZD034).

### References

1. China Internet Information Center (2017) A statistical report on the development of China. China's Internet Network [EB/OL].
2. Judicial Case Research Institute of the judiciary big data research institute (2017) Intellectual property rights infringement of the special judicial Report [EB/OL].
3. Supreme people's court of the People's Republic of China (2018) Notification on Bringing the Judicial Function into Full Play to Create a Good Legal Environment for Entrepreneurs' Innovation and Entrepreneurship [EB/OL].
4. General Office of the State Council, General Office of the Central Committee of the Communist Party of China. Opinions on Strengthening Reform and Innovation in the Field of Intellectual Property Trial [EB/OL].
5. Guanghai Lin (2018) Returning to the private law standard of intellectual property and giving play to the leading role of judicial protection [N]. Newspaper of the people's court.
6. Infringement of intellectual property rights.
7. Ministry of Public Security of the People's Republic of China (2011) Several issues concerning the application of law in handling criminal cases involving infringement of intellectual property rights [EB/OL].
8. Ministry of Public Security of the People's Republic of China (2017) Provisions of the Ministry of public security of the Supreme People's Procuratorate on handling economic crime cases by public security organs [EB/OL].
9. Ministry of Justice of the People's Republic of China (2016) General rules of judicial authentication procedure [EB/OL].
10. Judicial authentication administration of the Ministry of justice of the People's Republic of China (2014) Implementation specification for software similarity identification [EB/OL].
11. China Electronics Society (2015) Project of compiling and applying entries for science and Technology Encyclopedia in China [EB/OL].
12. China Electronics Society (2016) Project of compiling and applying entries for science and Technology Encyclopedia in China [EB/OL].
13. China Electronics Society (2015) Project of compiling and applying entries for science and Technology Encyclopedia in China [EB/OL].
14. Liu Z (2008) Research on anti-reverse engineering technology [D]. Shandong University of Science and Technology.
15. Zhang S (2016) Research on software protection scheme based on shell technology [D]. Chengdu University of Technology.

