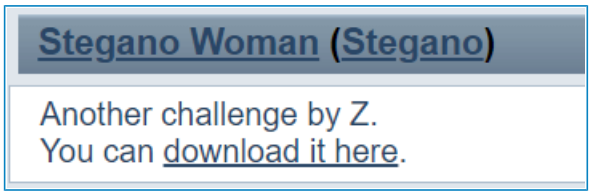
Stegano Woman (Stegano)答案 Write Up

原创

滕青山YYDS 于 2021-02-17 10:32:21 发布 107 收藏
文章标签： java

本文链接：https://blog.csdn.net/qq_34626094/article/details/113128295

分析

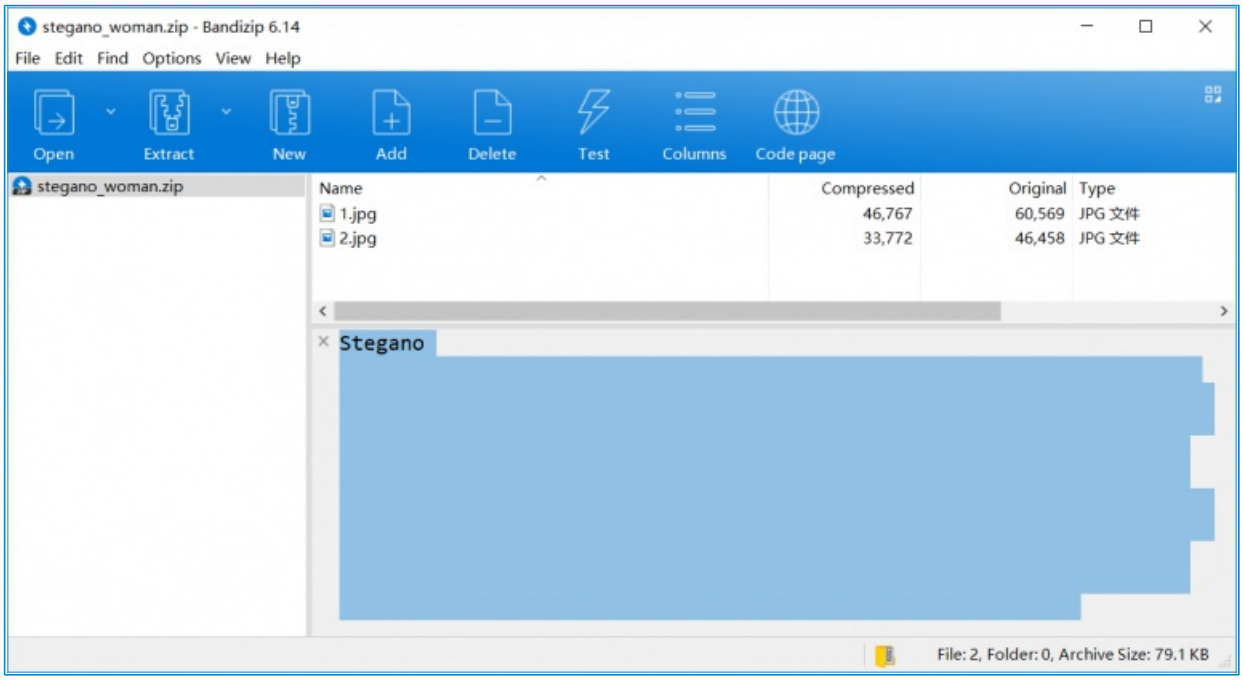


这道题目与图片内数据的隐写有一点点不同，它与 Training: Stegano I (Training, Stegano) 倒是比较像。实际上都是隐藏在二进制数据中的隐写。

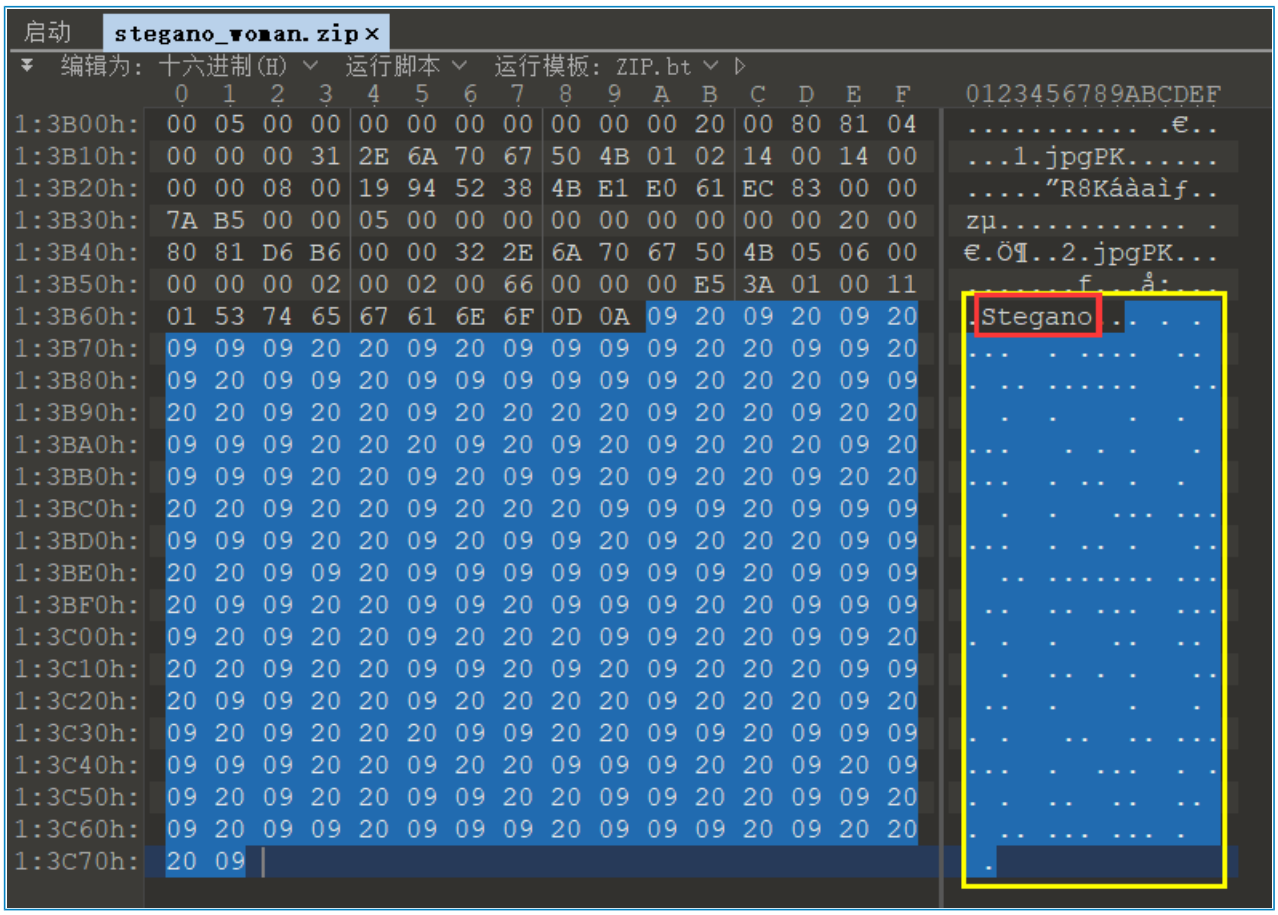
把题目的附件下载下来，直接解压会发现有两张图片。其实有两张图片的隐写还有别的方法，比如合并两张图片之类的。这道题的坑爹地方就是想把人往这个方向误导。

但是实际上如果不解压，直接打开压缩文件会有怎样的效果呢？

下方选中的部分实际上是压缩文件的注释部分。



看图里面我选中的注释，Stegano后面是不是还有一大串不可见的字符？原来猫腻就在这里。马上用010 Editor打开拉到最后看一下。



原来那一堆不可见的字符是两种字符组成的，09(tab)，20(space)。所以很容易想到这是使用两种不同字符的编码方式。

思考一下有什么编码方式是这样子的：二进制ASCII、Morse码、Bacon码……我大概就想到这么多。

下面我们来分析一下为什么后面两者是几乎不可能的。

Morse码，大家应该都知道它是由.-.-…——这种东西组成的。而且要搞清楚Morse码、Huffman编码、前缀编码的概念。Morse码和Huffman编码都是通过二叉树进行构造的。但是它们有一个很重要的区别——学过相关课程的人应该知道，Huffman编码全部采用了叶子节点。这样有什么好处？这样使得Huffman编码是前缀编码，即对字符集进行编码时，要求字符集中任一字符的编码都不是其它字符的编码的前缀。

再简单点，举个例子：如果001代表a，那么不存在一个字符会编码为0010或者0011。这就使得译码的时候，从头到尾看01串，看到到符合编码表中某一个字符的01子串就可以直接将其翻译为对应字符，而不需要顾虑有没有别的字符对应的子串的前缀与这个01子串相同。这样的编码方式虽然不同字符编码为不同长度的码，但是不需要分隔符就可以译码，而不会产生二义。

刚刚说Morse不一样，没有全部使用叶子节点，也就是说它肯定不是前缀编码。而且又因为它每个字符编码为不同长度的“嘀嗒串”，因此它必须要分隔符才能无二义地译码。看看密文，不像有分隔符的影子，因此这样直接去按Morse译码肯定有N种情况，基本不可能。

第二个Bacon码不是基本不可能，是根本不可能。为什么呢？Bacon比较多是用不同字体大小，不同字体等比较难发现的方式去加密的，这样使用比较不符合Bacon的使用场景。不过这只是一个“基本不可能”的因素。根本不可能是因为Bacon一个字符用长度为5的AB串表示。但是上面那一串东西长度是264，不是5的倍数，所以否定Bacon。

所以肯定是ASCII了。ASCII可以用7位也可以用8位的。算一下264不是7的倍数而是8的倍数（264=33*8），所以只能是8位编码了。又因为8位编码的时候第一位肯定肯定是0，所以显然“09”代表的是0，“20”代表的是1.

所以只要替换一下，转为ASCII码，再转为字符就可以得到Flag。

提示：16进制文本中`Stegano`后面的`0d0a`代表回车换行。所以不用管。

将其转换为二进制，并转换为ascii字符的**python2**代码如下：

```
#coding: utf-8
with open('stegano_woman.zip','rb') as f:
    text = f.read()

index = text.find('Stegano')
# 上面这个获取的是 S 的下标

text = text[index+9:]
#['S', 't', 'e', 'g', 'a', 'n', 'o', '\r', '\n']
# 所以这里要从 9 开始

text = ''.join('1' if x==' ' else '0' for x in text)

s=""
for i in range(len(text)/8):
   s+= chr(int(text[i*8:i*8+8],2))

print s
```

Python3代码：

```
with open('stegano_woman.zip','rb') as f:
    text = f.read()

index = text.find(bytes('Stegano','ascii'))
# 上面这个获取的是 S 的下标

text=text[index+9:]
#['S', 't', 'e', 'g', 'a', 'n', 'o', '\r', '\n']
# 所以这里要从 9 开始

text = ''.join('1' if x==32 else '0' for x in text)
s=""
for i in range(len(text)//8):
   s+= chr(int(text[i*8:i*8+8],2))

print (s)
```

解决

The solution is "dangerous life".
提交`dangerous life`即可。

参考：

[WeChall] Stegano Woman (Stegano) – Chiang E's Blog

wechall writeup — Evil