(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 106549948 A
(43)申请公布日 2017.03.29

(21)申请号 201610914988.1

(22)申请日 2016.10.20

(71)申请人 公安部第三研究所
地址 200031 上海市徐汇区岳阳路76号

(72)发明人 吴松洋 张旭 刘善军 刘欣

(74)专利代理机构 上海天翔知识产权代理有限公司 31224
代理人 刘常宝

(51)Int.Cl.
*H04L 29/06*(2006.01)
*H04L 12/58*(2006.01)
*G06F 21/62*(2013.01)

权利要求书2页 说明书6页

(54)发明名称

android平台下telegram应用多媒体取证方法

(57)摘要

本发明公开了一种android平台下telegram应用多媒体取证方法，其通过运行telegram应用的android平台下名称为cache4.db的文件中的media_v2表中存储的数据信息，进行android平台下telegram应用的多媒体信息的取证。本发明提供的取证方法有效实现在不破坏即时通信消息完整性和加密性的前提下，提取android平台下telegram应用的多媒体数据信息。且该取证方法可靠性和可操作性强。

CN 106549948 A

1.一种android平台下telegram应用多媒体取证方法,其特征在于,所述取证方法通过运行telegram应用的android平台下名称为cache4.db的文件中的media_v2表中存储的数据信息,进行android平台下telegram应用的多媒体信息的取证。

2.根据权利要求1所述的一种android平台下telegram应用多媒体取证方法,其特征在于,所述取证方法中首先获取运行telegram应用的android平台的ROOT权限;接着查找运行telegram应用的android平台下名称为cache4.db的文件;接着,将查找到的cache4.db文件导出到本地的文件系统中。

3.根据权利要求1所述的一种android平台下telegram应用多媒体取证方法,其特征在于,所述取证方法通过SQLite数据库查看工具从cache4.db文件中获取media_v2表。

4.根据权利要求1所述的一种android平台下telegram应用多媒体取证方法,其特征在于,所述取证方法针对不同多媒体信息的类型,由media_v2表获取cache4.db文件下message表中对应的message消息记录;再根据该message消息记录内容确定相应类型多媒体信息的特征字节及文件名格式信息,并由此针对相应类型多媒体信息进行取证。

5.根据权利要求4所述的一种android平台下telegram应用多媒体取证方法,其特征在于,所述取证方法中利用{0xc7,0xac,0x64,0x96}这四个字节确定音频文件中特征字节及文件名格式信息,并由此对音频文件进行取证。

6.根据权利要求5所述的一种android平台下telegram应用多媒体取证方法,其特征在于,对音频文件的取证过程如下:

(1)获取media_v2表中type字段为2的信息记录中的mid字段值;

(2)在cache4.db文件中的message表中查找到与media_v2表中mid值相同的message消息记录;

(3)将查找到的message消息记录的data字段转化为十六进制byte数组;

(4)在data字段的十六进制byte数组中寻找到以{0xc7,0xac,0x64,0x96}为逆序存放的第一次出现标识字节的开始位置,以该位置为游标基点,正向向后读取8个字节的数据,将该8个字节的数据逆序为一个数组,将这个8字节十六进制数据转化为10进制long类型整型数值,该值作为音频文件的dc_id值;

(5)将游标位置移动到data数组末尾处,以该处为基点,反向向前再读取8个字节的数据,将该8个字节的数据逆序为一个数组,将这个8字节十六进制数据转化为10进制long类型整型数值,该值作为音频文件的id值;

(6)在android平台的音频文件存储目录下遍历文件名称结构为dc_id+"_"+id+".ogg"格式的音频文件,完成音频文件取证。

7.根据权利要求4所述的一种android平台下telegram应用多媒体取证方法,其特征在于,所述取证方法中利用{0x53,0xd6,0x90,0x76}这四个字节确定图片文件中特征字节及文件名格式信息,并由此对图片文件进行取证。

8.根据权利要求7所述的一种android平台下telegram应用多媒体取证方法,其特征在于,对图片文件的取证过程如下:

(1)获取media_v2表中type字段为0的信息记录中的mid字段值;

(2)在cache4.db文件中的message表中查找到与media_v2表中mid值相同的message消息记录;

(3)将查找到的message消息记录的data字段转化为十六进制byte数组；

(4)在data字段的十六进制byte数组中寻找到以{0x53,0xd6,0x90,0x76}为逆序存放的最后一次出现标识字节的开始位置，以该位置为游标基点，正向向后读取4个字节的数据，将该4个字节的数据逆序为一个数组，将这个4字节十六进制数据转化为10进制long类型整型数值，该值作为图片文件的volume_id值；

(5)将游标位置以当前位置为基点，正向向后继续再读取4个字节的数据，将该4个字节的数据逆序为一个数组，将这个4字节十六进制数据转化为10进制long类型整型数值，该值作为图片文件的local_id值；

(6)在android平台的图片文件存储目录下遍历文件名称结构为volume_id+"_"+local_id+".jpg"或者其他类型的图片文件，完成图片文件取证。

9.根据权利要求4所述的一种android平台下telegram应用多媒体取证方法，其特征在于，所述取证方法中利用{0x38,0x8f,0xa3,0x91}这四个字节确定视频文件中特征字节及文件名格式信息，并由此对视频文件进行取证。

10.根据权利要求9所述的一种android平台下telegram应用多媒体取证方法，其特征在于，对视频文件的取证过程如下：

(1)获取media_v2表中type字段为0的信息记录中的mid字段值；

(2)在cache4.db文件中的message表中查找到与media_v2表中mid值相同的message消息记录；

(3)将查找到的message消息记录的data字段转化为十六进制byte数组；

(4)在data字段的十六进制byte数组中寻找到以{0x38,0x8f,0xa3,0x91}为逆序存放的第一次出现标识字节的开始位置，以该位置为游标基点，正向向后读取8个字节的数据，将该8个字节的数据逆序为一个数组，将这个8字节十六进制数据转化为10进制long类型整型数值，该值作为视频文件的id值；

(5)将游标位置移动到data数组末尾处，以该处为基点，反向向前再读取16个字节的数据，将该16个字节的数据逆序为一个数组，将这个16字节十六进制数据转化为10进制long类型整型数值，该值作为视频文件的dc_id值；

(6)在android平台的视频文件存储目录下遍历文件名称结构为dc_id+"_"+id+".mp4"的视频文件，完成视频文件取证。

# android平台下telegram应用多媒体取证方法

## 技术领域

[0001] 本发明涉及电子数据取证，特别涉及android平台下telegram应用多媒体取证技术。

## 背景技术

[0002] Telegram，中文名称为电报，是一个基于云的即时通讯服务应用。Telegram应用针对智能手机终端的不同，分别在iOS平台和android平台下有不同的手机应用。用户可以通过telegram应用发送任何类型的消息和交换照片、视频以及音频文件。电报账户绑定的是用户的电话号码，并通过短信或电话进行验证。用户可以将多台设备添加到自己的账户，并在每一个接收消息。用户连接的设备可单独或一次全部被删除。相关联的号码可以在任何时候被改变，用户的联系人将自动接收新的号码。用户可以设置一个别名，使他们能够发送和接收消息，而不会暴露他们的电话号码。

[0003] Telegram应用使用的登录身份验证的默认方法是基于SMS的单因素认证。Telegram的消息是基于云计算的消息，并可以在任何用户的连接设备的访问。用户可以共享照片，视频，音频信息和其它文件。用户可以单独或多达5,000个成员的组发送消息给其他用户。Telegram的传输以电报使者LLP服务器上的邮件进行加密，该服务为MTProto协议。根据消息报文传输的隐私性，所有数据都存储大量加密，并在每种情况下，加密密钥存储在不同的司法管辖区的其他几个DC上。这样使得取证工程师或物理入侵者无法获得访问用户数据。

## 发明内容

[0004] 针对现有Telegram应用的特点，需要一种针对telegram应用的电子数据取证方法。

[0005] 由此，本发明所要解决的技术问题是提供一种android平台下telegram应用多媒体取证方法，该取证方法能够在不破坏即时通信消息完整性和加密性的前提下提取相应的数据信息。

[0006] 为了达到上述目的，本发明提供的一种android平台下telegram应用多媒体取证方法，所述取证方法通过运行telegram应用的android平台下名称为cache4.db的文件中的media_v2表中存储的数据信息，进行android平台下telegram应用的多媒体信息的取证。

[0007] 优选的，所述取证方法中首先获取运行telegram应用的android平台的ROOT权限；接着查找运行telegram应用的android平台下名称为cache4.db的文件；接着，将查找到的cache4.db文件导出到本地的文件系统中。

[0008] 优选的，所述取证方法通过SQLite数据库查看工具从cache4.db文件中获取media_v2表。

[0009] 优选的，所述取证方法针对不同多媒体信息的类型，由media_v2表获取cache4.db文件下message表中对应的message消息记录；再根据该message消息记录内容确定相应类

型多媒体信息的特征字节及文件名格式信息，并由此针对相应类型多媒体信息进行取证。

[0010] 优选的，所述取证方法中利用{0xc7,0xac,0x64,0x96}这四个字节确定音频文件中特征字节及文件名格式信息，并由此对音频文件进行取证。

[0011] 优选的，对音频文件的取证过程如下：

[0012] (1)获取media_v2表中type字段为2的信息记录中的mid字段值；

[0013] (2)在cache4.db文件中的message表中查找到与media_v2表中mid值相同的message消息记录；

[0014] (3)将查找到的message消息记录的data字段转化为十六进制byte数组；

[0015] (4)在data字段的十六进制byte数组中寻找到以{0xc7,0xac,0x64,0x96}为逆序存放的第一次出现标识字节的开始位置，以该位置为游标基点，正向向后读取8个字节的数据，将该8个字节的数据逆序为一个数组，将这个8字节十六进制数据转化为10进制long类型整型数值，该值作为音频文件的dc_id值；

[0016] (5)将游标位置移动到data数组末尾处，以该处为基点，反向向前再读取8个字节的数据，将该8个字节的数据逆序为一个数组，将这个8字节十六进制数据转化为10进制long类型整型数值，该值作为音频文件的id值；

[0017] (6)在android平台的音频文件存储目录下遍历文件名称结构为dc_id+"_"+id+".ogg"格式的音频文件，完成音频文件取证。

[0018] 优选的，所述取证方法中利用{0x53,0xd6,0x90,0x76}这四个字节确定图片文件中特征字节及文件名格式信息，并由此对图片文件进行取证。

[0019] 优选的，对图片文件的取证过程如下：

[0020] (1)获取media_v2表中type字段为0的信息记录中的mid字段值；

[0021] (2)在cache4.db文件中的message表中查找到与media_v2表中mid值相同的message消息记录；

[0022] (3)将查找到的message消息记录的data字段转化为十六进制byte数组；

[0023] (4)在data字段的十六进制byte数组中寻找到以{0x53,0xd6,0x90,0x76}为逆序存放的最后一次出现标识字节的开始位置，以该位置为游标基点，正向向后读取4个字节的数据，将该4个字节的数据逆序为一个数组，将这个4字节十六进制数据转化为10进制long类型整型数值，该值作为图片文件的volume_id值；

[0024] (5)将游标位置以当前位置为基点，正向向后继续再读取4个字节的数据，将该4个字节的数据逆序为一个数组，将这个4字节十六进制数据转化为10进制long类型整型数值，该值作为图片文件的local_id值；

[0025] (6)在android平台的图片文件存储目录下遍历文件名称结构为volume_id+"_"+local_id+".jpg"或者其他类型的图片文件，完成图片文件取证。

[0026] 优选的，所述取证方法中利用{0x38,0x8f,0xa3,0x91}这四个字节确定视频文件中特征字节及文件名格式信息，并由此对视频文件进行取证。

[0027] 优选的，对视频文件的取证过程如下：

[0028] (1)获取media_v2表中type字段为0的信息记录中的mid字段值；

[0029] (2)在cache4.db文件中的message表中查找到与media_v2表中mid值相同的message消息记录；

[0030] (3)将查找到的message消息记录的data字段转化为十六进制byte数组；

[0031] (4)在data字段的十六进制byte数组中寻找到以{0x38,0x8f,0xa3,0x91}为逆序存放的第一次出现标识字节的开始位置，以该位置为游标基点，正向向后读取8个字节的数据，将该8个字节的数据逆序为一个数组，将这个8字节十六进制数据转化为10进制long类型整型数值，该值作为视频文件的id值；

[0032] (5)将游标位置移动到data数组末尾处，以该处为基点，反向向前再读取16个字节的数据，将该16个字节的数据逆序为一个数组，将这个16字节十六进制数据转化为10进制long类型整型数值，该值作为视频文件的dc_id值；

[0033] (6)在android平台的视频文件存储目录下遍历文件名称结构为dc_id+"_"+id+".mp4"的视频文件，完成视频文件取证。

[0034] 本发明提供的取证方法有效实现在不破坏即时通信消息完整性和加密性的前提下，提取android平台下telegram应用的多媒体数据信息。且该取证方法可靠性和可操作性强。

## 具体实施方式

[0035] 为了使本发明实现的技术手段、创作特征、达成目的与功效易于明白了解，下面结合具体实例，进一步阐述本发明。

[0036] 本发明提供的取证方法基于运行telegram应用的android平台下名称为cache4.db的文件中的media_v2表中存储的数据信息，进行android平台下telegram应用的多媒体信息的取证。

[0037] 具体为针对不同多媒体信息的类型，由media_v2表中的数据信息获取cache4.db文件下message表中对应的message消息记录；再根据该message消息记录内容确定相应类型多媒体信息的特征字节及文件名格式信息，并由此针对相应类型多媒体信息进行取证。

[0038] 基于该原理，以对android智能手机中telegram应用的多媒体数据进行取证为例，对本方案进行具体说明。

[0039] 整个取证过程主要分为取证前准备工作以及取证操作。

[0040] (一)取证前准备工作，获取cache4.db文件以及media_v2表。

[0041] 1.首先获取对android手机的root权限。可以采用对手机进行root的方法获取该手机的root权限，也可以采用临时root的方法获取该手机的root权限。在获取了该手机root权限的前提下，通过使用adb命令，获取android手机文件系统/data/data/org.telegram.messenger/files路径下的所有文件列表，遍历该路径下的所有文件，查找一个名称为cache4.db的文件，该cache4.db文件为SQLite数据库文件，分析该文件可以找出许多有关多媒体存储位置的信息。再次通过adb的pull命令，将该cache4.db文件从android手机的文件系统导出到本地的文件系统中单独分析。

[0042] 2.针对已导出到本地文件系统中的cache4.db文件，通过SQLite数据库查看工具直接找开，可以发现该SQLite数据库中包含多张数据表，针对不同的表可以取证出不同的telegram应用数据信息。

[0043] users表可以分析出telegram应用的用户账户信息；

[0044] contacts表、dialogs表可以分析telegram应用的好友联系人信息；

[0045] dialogs表可以获取当前与手机账户最近有聊天记录的好友，综合contacts表和最近的聊天好友信息合并成最终的好友联系人列表；

[0046] message表包含telegram应用所有好友的聊天记录信息；

[0047] media_v2表可以分析telegram应用的多媒体信息。

[0048] (二)通过media_v2表中存储的数据信息对多媒体信息进行取证。

[0049] (1)android平台下telegram应用的多媒体数据一般都会存储在android手机的手机存储卡上，体现在android文件系统中就是在/sdcard路径下，该sdcard路径下存储了android各类应用的多媒体或者其他不宜直接存储在手机内存中的数据文件，在sdcard路径下可以找到一个名为Telegram的目录，该目录即为telegram应用的多媒体数据存储目录，在该目录下针对不同的多媒体类型又存在多个目录。例如图片文件存储在Telegram Images目录下，音频文件存储在Telegram Audio目录下，视频文件存储在Telegram Video目录下。

[0050] 由此在进行常规取证时，可以将Telegram目录下不同类型的多媒体文件通过adb命令全部导出到本地文件系统中。

[0051] (2)media_v2表中的每一条记录即代表一条多媒体信息记录。申请人通过大量研究cache4.db文件中的多张表，发现message表中的mid字段和media_v2表中的mid字段存在关联关系。message表中的mid字段与media_v2表中的mid字段含义相同，都是该条多媒体消息文件的唯一可区分消息标识。此时通过media_v2表中的type字段可以区分该条多媒体消息的类型，当type值为2时，表明该多媒体文件为语音文件；当type值为0时，表明该媒体文件为图片或者视频文件。

[0052] (3)针对media_v2表中的type字段为2的多媒体文件，申请人通过大量研究发现在sdcard下的Telegram/Telegram Audio目录中存储着后缀名为ogg的音频文件，ogg是一种特殊的音频文件，通过解码处理可转化为能正常播放的音频文件。media_v2表通过mid字段与message表相关联，在message表中查找到与media_v2表中mid值相同的message消息记录，分析message表中的data字段，发现data字段的表结构类型为二进制byte数组类型。

[0053] 将该二进制byte数组转化为十六进制byte数组进行分析，发现所有的音频文件data数组中必然存在{0xc7,0xac,0x64,0x96}4个标识字节。将这4个标识字节定义为音频文件的开始标识字节，这4个标识字节在data数组总是以逆序的方式进行存储。故可确定多媒体音频文件为多字节数据存储文件，而且多字节数据存放为逆序存放。

[0054] 多字节数据存放顺序与CPU有关，微处理器中的存放顺序有正序(Big-Endian)和逆序(Little-Endian)，也称为大端存储和小端存储。例如常见的Intel系列使用的编码方式属于逆序存放；某些RISC架构的CPU，如IBM的Power-PC等属于正序存放。两种编码区别在于：正序(Big-Endian)高位字节存入低地址，低位字节存入高地址，依次排列。逆序(Little-Endian)低位字节存入低地址，高位字节存入高地址，反序排列。

[0055] (4)针对media_v2表中的type字段为0的多媒体文件，申请人通过大量研究发现在sdcard下的Telegram/Telegram Image目录中存储着后缀名为jpg或者其他类型后缀的图片文件。media_v2表通过mid字段与message表相关联，在message表中查找到与media_v2表中mid值相同的message消息记录，分析message表中的data字段，发现data字段的表结构类型为二进制byte数组类型。

[0056] 将该二进制byte数组转化为十六进制byte数组进行分析，发现所有的图片文件data数组中必然存在{0x53,0xd6,0x90,0x76}4个标识字节。将这4个标识字节定义为图片文件的开始标识字节，这4个标识字节在data数组总是以逆序的方式进行存储。故可确定多媒体音频文件为多字节数据存储文件，而且多字节数据存放为逆序存放。

[0057] (5)针对media_v2表中的type字段为0的多媒体文件，申请人通过大量研究发现在sdcard下的Telegram/Telegram Video目录中存储着后缀名为mp4的视频文件。media_v2表通过mid字段与message表相关联，在message表中查找到与media_v2表中mid值相同的message消息记录，分析message表中的data字段，发现data字段的表结构类型为二进制byte数组类型。

[0058] 将该二进制byte数组转化为十六进制byte数组进行分析，发现所有的视频文件data数组中必然存在{0x38,0x8f,0xa3,0x91}4个标识字节。将这4个标识字节定义为图片文件的开始标识字节，这4个标识字节在data数组总是以逆序的方式进行存储。故可确定多媒体音频文件为多字节数据存储文件，而且多字节数据存放为逆序存放。

[0059] 基于上述的原理，本实例中进行对音频文件、图片文件以及视频文件进行取证的过程分别如下。

[0060] 1、对音频文件的取证过程如下：

[0061] (1)获取media_v2表中type字段为2的信息记录中的mid字段值；

[0062] (2)在cache4.db文件中的message表中查找到与media_v2表中mid值相同的message消息记录；

[0063] (3)将查找到的message消息记录的data字段转化为十六进制byte数组；

[0064] (4)在data字段的十六进制byte数组中寻找到以{0xc7,0xac,0x64,0x96}为逆序存放的第一次出现标识字节的开始位置，以该位置为游标基点，正向向后读取8个字节的数据，将该8个字节的数据逆序为一个数组，将这个8字节十六进制数据转化为10进制long类型整型数值，该值作为音频文件的dc_id值；

[0065] (5)将游标位置移动到data数组末尾处，以该处为基点，反向向前再读取8个字节的数据，将该8个字节的数据逆序为一个数组，将这个8字节十六进制数据转化为10进制long类型整型数值，该值作为音频文件的id值；

[0066] (6)当获取到了音频文件的dc_id和id值之后，在sdcard的Telegram/Telegram Audio目录下遍历文件名称结构为dc_id+"_"+id+".ogg"格式的音频文件，由此在Telegram Audio目录下会取证到该音频文件。

[0067] 2、对图片文件的取证过程如下：

[0068] (1)获取media_v2表中type字段为0的信息记录中的mid字段值；

[0069] (2)在cache4.db文件中的message表中查找到与media_v2表中mid值相同的message消息记录；

[0070] (3)将查找到的message消息记录的data字段转化为十六进制byte数组；

[0071] (4)在data字段的十六进制byte数组中寻找到以{0x53,0xd6,0x90,0x76}为逆序存放的最后一次出现标识字节的开始位置，以该位置为游标基点，正向向后读取4个字节的数据，将该4个字节的数据逆序为一个数组，将这个4字节十六进制数据转化为10进制long类型整型数值，该值作为图片文件的volume_id值；

[0072] (5)将游标位置以当前位置为基点，正向向后继续再读取4个字节的数据，将该4个字节的数据逆序为一个数组，将这个4字节十六进制数据转化为10进制long类型整型数值，该值作为图片文件的local_id值；

[0073] (6)当获取到了图片文件的volume_id和local_id值之后，在sdcard的Telegram/Telegram Image目录下遍历文件名称结构为volume_id+"_"+local_id+".jpg"或者其他类型的图片文件，由此在Telegram Image目录下会取证到该图片文件。

[0074] 3、对视频文件的取证过程如下：

[0075] (1)获取media_v2表中type字段为0的信息记录中的mid字段值；

[0076] (2)在cache4.db文件中的message表中查找到与media_v2表中mid值相同的message消息记录；

[0077] (3)将查找到的message消息记录的data字段转化为十六进制byte数组；

[0078] (4)在data字段的十六进制byte数组中寻找到以{0x38,0x8f,0xa3,0x91}为逆序存放的第一次出现标识字节的开始位置，以该位置为游标基点，正向向后读取8个字节的数据，将该8个字节的数据逆序为一个数组，将这个8字节十六进制数据转化为10进制long类型整型数值，该值作为视频文件的id值；

[0079] (5)将游标位置移动到data数组末尾处，以该处为基点，反向向前再读取16个字节的数据，将该16个字节的数据逆序为一个数组，将这个16字节十六进制数据转化为10进制long类型整型数值，该值作为视频文件的dc_id值；

[0080] (6)当获取到了音频文件的id和dc_id值之后，在sdcard的Telegram/Telegram Video目录下遍历文件名称结构为dc_id+"_"+id+".mp4"的视频文件，由此在Telegram Video目录下会取证到该视频文件。

[0081] 以上显示和描述了本发明的基本原理、主要特征和本发明的优点。本行业的技术人员应该了解，本发明不受上述实施例的限制，上述实施例和说明书中描述的只是说明本发明的原理，在不脱离本发明精神和范围的前提下，本发明还会有各种变化和改进，这些变化和改进都落入要求保护的本发明范围内。本发明要求保护范围由所附的权利要求书及其等效物界定。