C23 Diving Into Blockchains Contents: The Bitcoin Snapshot

Oliver Giudice, PhD, Roma 00044, ITALY*

Learning Overview: The goal of this presentation is to present a method for blockchain analysis with the goal of discovering non-transactional information, such as image files, text files, etc. Attendees will be introduced to the basics of blockchain technologies and Application Programming Interfaces (APIs) of the bitcoin protocol to understand end-user capabilities. Finally, blockchain diving techniques will be presented to assist attendees in creating automatic tools for blockchain analysis and content retrieval.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by demonstrating how blockchain technology is a novel field in which new crimes are emerging (illicit content sharing, Intellectual Property (IP) violations, etc.), thus highlighting the need for new methods, techniques, and expertise.

Blockchain technology in recent years has been defined as the new generation of internet, or rather a type of internet of transactions. There is no definition of blockchain that is universal, but it is a concept that has several more or less valid interpretations.

Blockchain technology actually belongs to an even broader concept, Distributor Ledger Technology (DLT). DLTs are distributed databases, meaning they can be updated, managed, controlled, and coordinated not only at the central level, but in a distributed way, by all the actors participating in the DLT. Each user can manage a node, but each transaction must first be approved by most participants in the network. DLTs use independent computers (called nodes) to record, share, and synchronize transactions in the corresponding ledgers (master books). The use of new technologies could certainly have a major impact on the world economy due to many elements, such as security, the speed of transmission of transactions, and anonymity. DLT applications could very quickly replace many tasks that are still being performed manually today.

Bitcoin is an implementation of blockchain technology, which is the most known and used today. However, while blockchains can be a great technology for transactions of value, they do not contain data related only to transactions. However, they can contain other information, such as text files or images.

The problem arises when this data is illegal material; anyone who memorized and shared the material could be accused of possession of that illegal information. Thus, the need for a blockchain analytics tool arises to detect non-transactional content diving into a large-scale quantity of data.

Writing non-transactional information is very easy; specifically in bitcoin, it is possible to insert in the place of the address any string, be it an address, a text, or a link. In doing so, of course, there is a waste of money, but irrelevant sums could be inserted. Once the desired string has been inserted, the shared material is visible and can be downloaded from every node present in the network.

As an example, it is possible to look at the third block of the bitcoin blockchain and obtain as a result a set of characters that make up the face of a man.

Techniques will be presented that are able to automatically detect non-transactional information in the bitcoin blockchain by means of address parsing, transactions content analysis, and analysis of the OP_RETURN field.

The simple technique that will be described in this presentation is the ability to identify more than 1,500 non-transactional contents present on the bitcoin blockchain until December 2017. All those contents can be divided into the following types: source code files, HTML files, image files, and text files.

Finally, issues related to illegal material on blockchains will be discussed with respect to contents extracted from the bitcoin blockchain.

Bitcoin Forensics, Bitcoin Content Analysis, Illicit Media Sharing