Indian Institute of Technology
Bombay

Internship Final Report

Blockchain Forensic Tool

Author:
Syed Saud Hasan

Supervisor:
Prof. G Sivakumar
Dr. Vishvas Patil

July 6, 2017

Acknowledgement

I would like to express my deepest appreciation to **Prof Madhu N. Belur** and **Dr. Kuntal Deka** who helped me in getting a intern project at IIT Bombay. A special gratitude I give to my intern guide **Prof. G Sivakamar** and co-guide **Dr. Vishvas Patil**, whose contribution in stimulating suggestions and encouragement, helped me to coordinate my project.

1 Abstract

Blockchain technology is being touted as next major disruptive technology in the field of computing world. It is going to revolutionize different sectors, specially the financial technology sector. During the internship I did survey of Blockchain technology, knew how it works, which companies are working in this domain, what are existing tools and how we can make a better one. Finally we came up with a Blockchain forensic tool that could make life of government agencies easy.
Apart from Blockchain technology I attended few workshops and met with people who working in this field.

2 Blockchain Forensic Tool

2.1 Intro - Bitcoin and Blockchain

Bitcoin is the first digital cryptocurrency and Blockchain is the underlying technology that supports Bitcoin. A Blockchain is a public ledger of all Bitcoin transactions that have ever been executed. It is constantly growing as new blocks are added to it with a new set of transactions. The blocks are added to the Blockchain in a linear, chronological order. Each node (computer connected to the Bitcoin network using a client that performs the task of validating and relaying transactions) gets a copy of the Blockchain, which gets downloaded automatically upon joining the Bitcoin network. The Blockchain has complete information about the addresses and their balances right from the genesis block to the most recently completed block.

2.2 Motivation

The legal status of Bitcoin varies substantially from country to country and is still undefined or changing in many of them. Whilst the majority of countries do not make the usage of Bitcoin itself illegal (with the exceptions of: Bangladesh, Bolivia, Ecuador and Kyrgyzstan), its status as money (or a commodity) varies, with differing regulatory implications. While some countries have explicitly allowed its use and trade, others have banned or restricted it. Likewise, various government agencies, departments, and courts have classified Bitcoins differently.

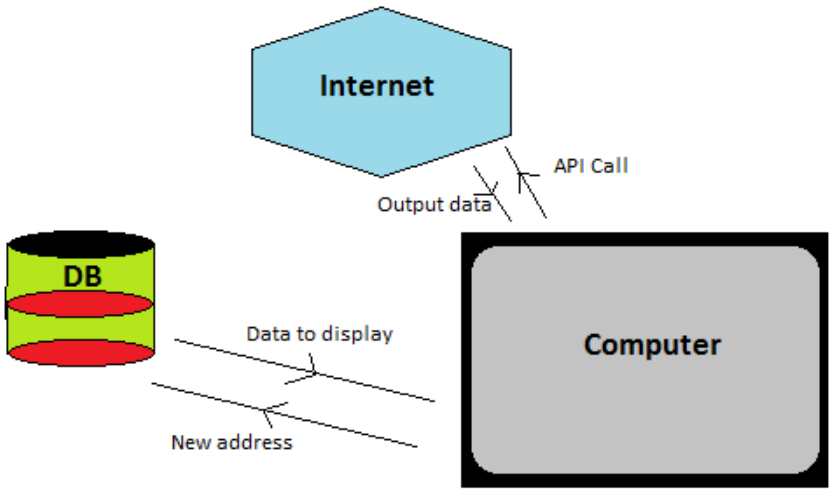


Figure 1: Architecture

The transactions in Blockchain are completely anonymous and hence can't be traced easily by the government agencies. Hence, our project aimed at designing a forensic tool for Blockchain technology which can help government and security agencies to take better insight of Bitcoin transaction.

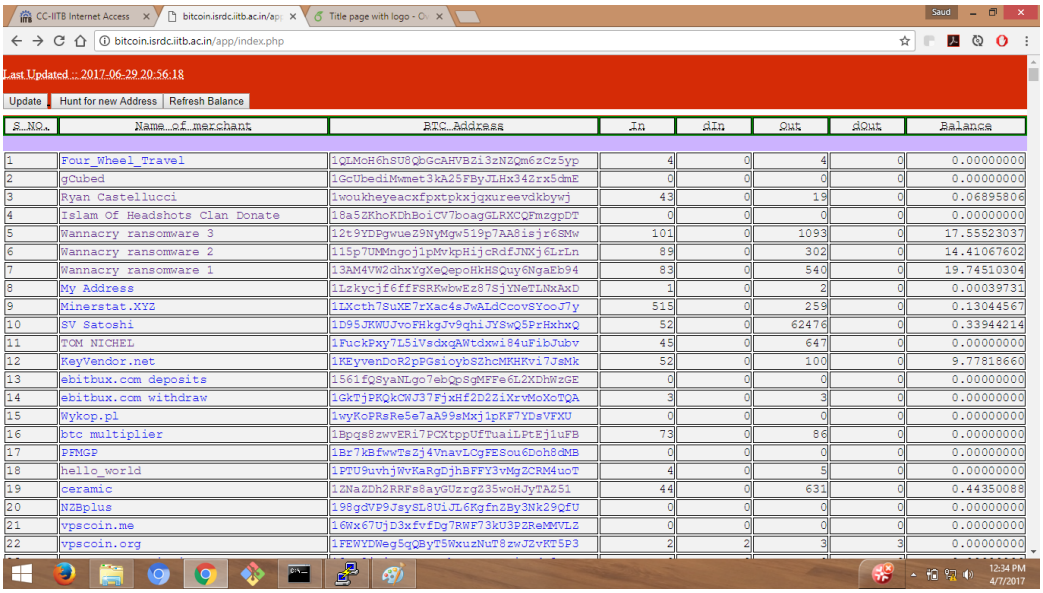
2.3 Approach

Core idea behind the tool is to club publicly available Blockchain ledger information with auxiliary information. Auxiliary information can be easily collected form different websites, blogs and social media sites using available api's and web crawler. These informations must be stored in some database.

For this tool I have used API of Blockchain.info. They provide mapping of address to their respective names and URL's. Usually they collect mapping of address to name by croud sourcing. All the Bitcoin address's related to concerned address is matched with the data available. During the representation of address in our tool if we know the name mapping of a particular address then the name is displayed instead of the address.

2.4 Model

To design this tool (Figure 1) we have used php, MySql, HTML, CSS, and python. Web crawler crawls the Internet and adds newly gathered address and other information in Mysql database. Scripts uses the some information from database and rest from the Internet to display information on browser.



S_NO.	Name_of_merchant	BTC_Address	In	dIn	Out	dOut	Balance
1	Four_Wheel_Travel	1QLMoH6hSU8QbGcAHVBZi3zNZQm6zCz5yp	4	0	4	0	0.00000000
2	gCubed	1GcUbediMwmet3kA25FByJLHx34Zrx5dmE	0	0	0	0	0.00000000
3	Ryan Castellucci	1woukheyeacxfpxtpkxjqxureevdkbywj	43	0	19	0	0.06895806
4	Islam Of Headshots Clan Donate	18a5ZKhoKDhBoiCV7boagGLRXCQFmzgpDT	0	0	0	0	0.00000000
5	Wannacry ransomware 3	12t9YDPgwueZ9NyMgw519p7AA8isjr6SMw	101	0	1093	0	17.55523037
6	Wannacry ransomware 2	115p7UMMngoj1pMvkpHijcRdfJNXj6LrLn	89	0	302	0	14.41067602
7	Wannacry ransomware 1	13AM4VW2dhxYgXeQepoHkHSQuy6NgaEb94	83	0	540	0	19.74510304
8	My Address	1Lzkycjf6ffFSRKwbwEz87SjYNeTLNxAxD	1	0	2	0	0.00039731
9	Minerstat.XYZ	1LXcth7SuXE7rXac4sJwALdCcovSYooJ7y	515	0	259	0	0.13044567
10	SV Satoshi	1D95JKWUJvoFHkgJv9qhiJYSwQ5PrHxhxQ	52	0	62476	0	0.33944214
11	TOM_NICHEL	1FuckPxy7L5iVsdxqAWtdxwi84uFibJubv	45	0	647	0	0.00000000
12	KeyVendor.net	1KEyvenDoR2pPGsioybSZhcMKHKvi7JsMk	52	0	100	0	9.77818660
13	ebitbux.com deposits	1561fQSyaNLgo7ebQpSgMFFe6L2XDhWzGE	0	0	0	0	0.00000000
14	ebitbux.com withdraw	1GkTjPEQkCWJ37FjxHf2D2ZiXrvMoXoTQA	3	0	3	0	0.00000000
15	Wykop.pl	1wyKoPRsRe5e7aA99sMxj1pKF7YDsVFXU	0	0	0	0	0.00000000
16	btc multiplier	1Bpqs8zwvERi7PCXtppUfTuaiLPtEj1uFB	73	0	86	0	0.00000000
17	PFMGP	1Br7kBfwwTsZj4VnavLCgFESou6Doh8dMB	0	0	0	0	0.00000000
18	hello_world	1PTU9uvhjWvKaRgDjhBFFY3vMgZCRM4uoT	4	0	5	0	0.00000000
19	ceramic	1ZNaZDh2RRFs8ayGUzrgZ35woHJyTAZ51	44	0	631	0	0.44350088
20	NZBplus	198gdVP9JsySL8UiJL6KgfnZBy3Nk29QfU	0	0	0	0	0.00000000
21	vpscoin.me	16Wx67UjD3xfvfDg7RWF73kU3PZReMMVLZ	0	0	0	0	0.00000000
22	vpscoin.org	1FEWYDWeg5qQByT5WxuzNuT8zwJZvRT5P3	2	2	3	3	0.00000000

Figure 2: index.php

2.5 Interfaces

This tool provides few interfaces which help user in understanding the relations between the address easily.

1. *index.php* (Figure 2) :: This page show details of all the address that we have in our database. clicking on a particular name leads the user to the home page where that address is advertised by the merchant, *"In"* represents number of address's which have given BTC to the corresponding address, similarly *"Out"* represents number of Bitcoin address's which have received BTC from corresponding address, *"dIn"* represents number of address's from whom corresponding address has received BTC's since last refresh, similarly *"dOut"* represents number of Bitcoin address's who have received BTC from corresponding address since last refresh, *"Balance"* represents the current balance of corresponding address. Clicking on any of the *"BTC Address"* leads user to *addrhistory.php* page.

2. *addrhistory.php* (Figure 3) :: This page also contains 3 column's. Address in the mid column is said to be key address, all address's in left column have given a particular number of BTC's to key address, similarly key address has given particular numbers of BTC's to address's mentioned in right column. Page summary at the footer of the page displays the total number of address's who have given BTC's to key address, number of such address's

Figure 3: addrhistory.php

which have received BTC's from key address and total fees that key address has given till now.

3. *filter.php* (Figure 4) :: This page contains several input fields corresponding to every filter. Each filter has different use case, like i). #Filter1- Given Bitcoin address and a range of BTC the output will be all those address which have done transaction in given range with the provided key address and it will be displayed on *displayfilter1.php*. ii). #Filter2- Input range of BTC and the output will be set of all address in our local database which have balance in certain range, output will be displayed on *displayfilter2.php*. This filter can be used to know the economic class of a particular Bitcoin user. iii). #Filter3- Input a range of Block height and a range of BTC value, the output will be set of all transactions in given range of blocks falling in certain range of BTC value, the output will be displayed in *displayfilter3.php*. #Filter3 can be used to know, how many people were affected by recent attack of *Petya ransomware*. Tool only needs to know that at what block height Petya started receiving ransom and what is the amount that they have asked which is around $300 or 0.0004BTC. Both of the informations are easily available in public domain. (Figure 5) shows Address's which have done transaction in range of 0.000395BTC to 0.000415BTc on 4/7/2017 i.e. from block height of 473593 to block height of 473756. This analysis may not be proving anything but can give a reason to someone who is interested in knowing about who were attacked.

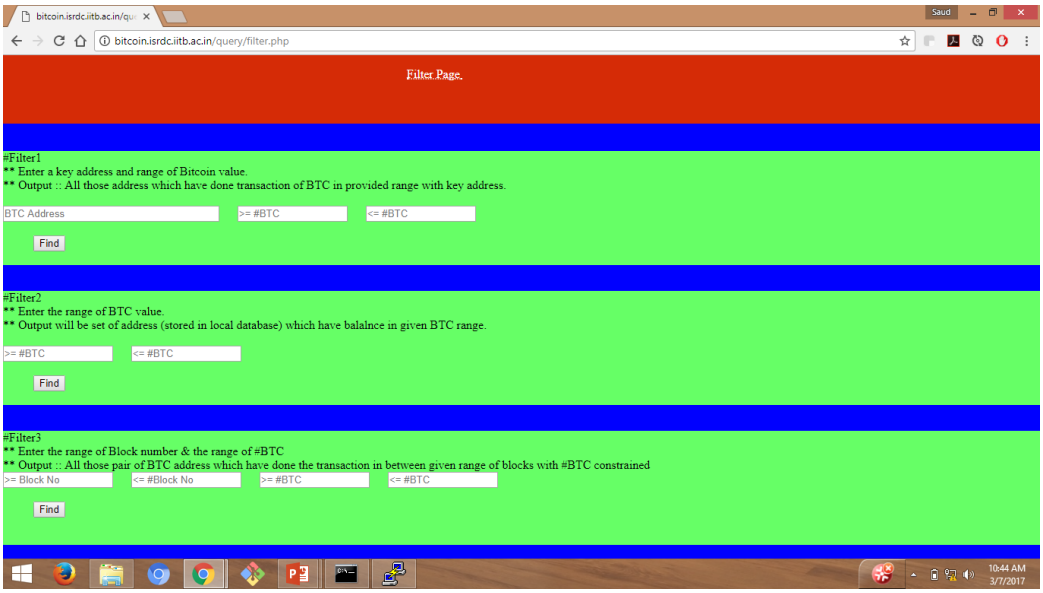


Figure 4: filter.php

4. Few buttons are also provided in different pages in the header section. i). *Update* :: Once this button is pressed script in the background will look for new registered address which have name associated with them, it will add them in local database and will do a complete refresh of the local database i.e. all the field corresponding to all address's in the local database will be updated. (Remember this will take a lot of time as we are querying to API of Blockchain.info for every block). ii). *Hunt for new address* :: Scrip5t behind this button will only add new address by searching them in next 1000 blocks. Initial block is saved in mysql database and is updated automatically after every time this button is pressed. iii). *Refresh Balanace* :: This button will only update all fields of every Bitcoin address present in our local database. iv). *Query Page* :: This page will let user to *filter.php* page.

2.6 Shortcomings

This forensic tool is designed using the API's of Blockchain.info which is a third party. This software can go down once Blockchain.info changes their API. Also to get better result and utilize this software fully this tool must have access to large amount of public data which is not usually available to general public.

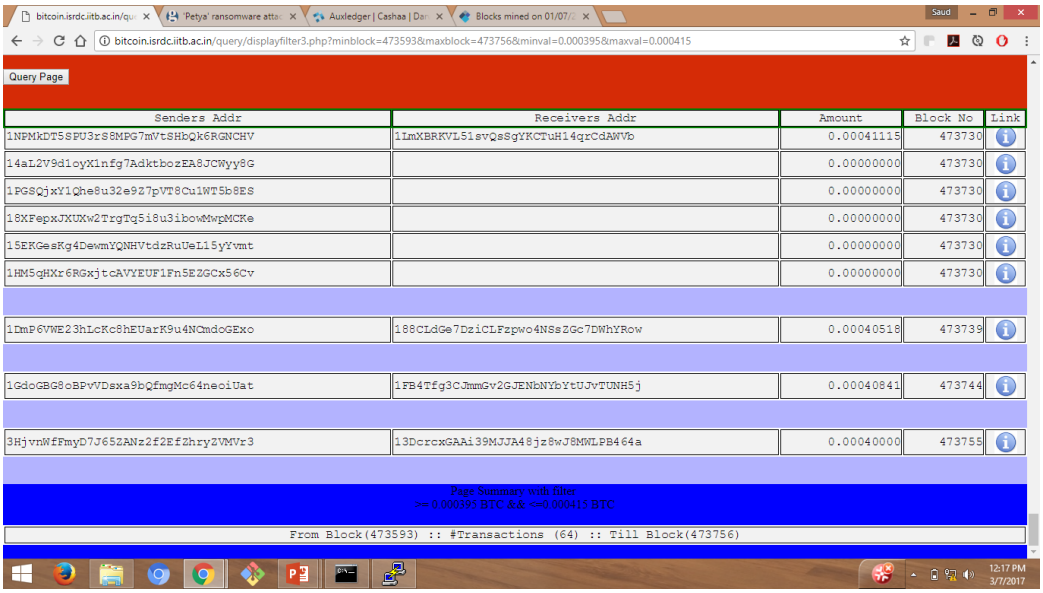


Figure 5: Result of filter3

2.7 Solutions

Above stated shortcomings can be easily solved by providing hardware support to this tool. A full node can be installed and ledger data can be fetched from there, and it will be reliable source also. In future if some change is required in the API then corresponding change can be done in scripts of the tools also. Huge amount of public data is easily accessible by government agencies hence our tool can work excellent with provided conditions.

3 Presentations, workshop and meetings

3.1 FOSSEE lab

PhD students of FOSSEE lab IITB explained about the projects the open source projects that are under pipeline. PYTHON TEXTBOOK COMPANIONS ON CLOUD was one of them, where a student can learn and give test of Python. Another open source project about which they talked was MIXED MODE SIMULATION IN ESIM using NgSpice and GHDL. In NGHDL, NgSpice is used to simulate the analog components and GHDL is used to simulate the digital components, where the analog and digital components are communicating through socket.

3.2 QIP Workshop- 2017

Helped PhD students and worked as TA for lab for this workshop. Attended a session by NanadKumar Sarvade (Chief Executive Officer, Reserve Bank Information Technology Pvt Ltd (ReBIT)) on security of financial institutions and Dr Vishvas Patil on Blockchain technology.

3.3 Latex Workshop

Attended workshop on Latex organized by Spoken Tutorial Project IITB.

3.4 Talk by Kumar Gaurav

Attended a talk by Kumar Gaurav on Blockchain technology. Kumar Gaurav is chaiman of Auxexis group and founder of caashaa. Auxexis group is among top 50 Blockchain companies of the world.

4 References

5.1 https://developers.coinbase.com/api/v2
5.2 https://blockchain.info/api